A Mixed-Method Response Surface Analysis of the OCM

Alexander Garnås

Master of Work- and Organizational Psychology

University of Oslo

May 2013

Acknowledgements

First I would like to thank my supervisor Roald Bjørklund for his advice and support. I would also like to thank Jon Anders Lone for his help with the data. Also, I would like to thank Trond Myklebust for helping me getting to know the police organization. And to my fellow students: thank you! It's been fun!

And of course, my parents. For all their support over the years.

Alexander Garnås, May 2013, Oslo

Table of Contents

Abstract

The aim of this study was to investigate if polynomial regression and response surface methodology could be used to study the relationship between statements from interviews and scores from questionnaires. The questionnaire used in the study was the Organizational Climate Measure (OCM). The study used a concurrent mixed model design and the qualitative and quantitative strands of data were equally emphasized. Content analysis was used to make the data from interviews comparable to the questionnaires. The polynomial regression analyses could account for the relationship between interviews and questionnaires on 10 of the 17 scales of the OCM. An examination of the response surfaces showed that the technique could be used to examine both convergent and divergent results between the two strands of data.

A Mixed-Method Response Surface Analysis of the OCM

How employees perceive their work environment has for a long time been a central question within the research of work- and organizational psychology. The shared perceptions of employees has been referred to as the organizational climate. According to Schneider (2013), organizational climate refers to "the shared perceptions of and the meaning attached to the policies, practices, and procedures employees experience and the behaviours they observe getting rewarded and that are supported and expected" (p. 362). The concept has spurred a lot of interest within research because employees will use their understanding of an organizations climate as a guideline for appropriate behaviour (Ployhart, Schmitt, & Schneider, 2006).

The climate dimensions suggested by these models have most commonly been measured using questionnaires with close-ended items (Patterson et al., 2005). This approach to gathering data offers many benefits, but suffers also from several important drawbacks. This has led some researchers to recommend the use of interviews as an additional tool to exploring the research area of interest (Bakker & Demerouti, 2007).

Close-ended questionnaires have the benefit of being fast for the respondent to complete and easy to interpret for the researcher. They also ensure a uniformity in the answers that simplifies later analysis. But Close-ended questionnaires prompts the respondent to asses categories that a model assumes are relevant to the employee. By prompting a respondent to relate to words and categories that he or she normally would not use, the results of close-ended questionnaires might not reflect the respondents actual experience. In open-ended interviews on the other hand, the interviewee is free to use his or her own words and categories. The results could more likely reflect the persons actual experience of his or her environment. Conducting interviews also poses some difficult challenges. Skinner (1957) referred to language use as verbal behaviour. He argued that the principles of conditioning also applied to how we communicate with others. The interviewer must therefore be mindful of his or her administration of reinforcement during interviews. The use of reinforcement and leading questions can influence the interviewee and distort the results. Interviews can also be time consuming to analyse.

Mixed-methods is a new approach to research that emphasizes the combined use of both qualitative and quantitative data. Mixed method researchers argue that this combination reduces the drawbacks of using either in isolation. To date there are few studies that have compared the use of interviews and questionnaires. This study will investigate the relationship

between data gathered from interviews and data gathered using the Organizational Climate Measure (Patterson et al., 2005).

Mixed-methods.

Mixed methods is a new and emerging approach to research (Plano Clark, 2010). It was formally established as a distinct methodological approach around the year 2000 (Teddlie & Tashakkori, 2010). The work of Campebell and Fiske (1959) can be said to have anticipated the advent of mixed methods. They argued that multiple methods should be used in the validation of psychological traits. But, as Creswell (2011) has pointed out, their use of the multitrait-multimethod matrix consisted of using multiple quantitative methods. The combined use of both qualitative and quantitative methods first appeared in the 1970s. Mixed methods has been described as a third methodological movement, besides the qualitative and quantitative approaches to research (Tashakkori & Teddlie, 2003). In mixed methods both qualitative and quantitative strands of data are combined. Mixed method researchers argue that the blending of these approaches reduces the drawbacks of using either in isolation. The benefits inherent in one approach will compensate for the weaknesses in the other.

Johnson and Onwuegbuzie (2007) have defined mixed methods research as "the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration". One of the core characteristics of the mixed method approach is the rejection of the incompatibility thesis. The incompatibility thesis claims that qualitative and quantitative methods cannot be combined because they are rooted in fundamentally different paradigms. While quantitative methods are rooted in positivism, qualitative methods are rooted in constructivism.

The organizational climate measure (OCM).

The organizational climate measure (OCM) is a global multidimensional measure of organizational climate. The measure consists of 82 items that assess 17 scales. The measure is has several features that sets it apart from other similar instruments.

The measure is theoretically grounded by being based on the Competing Values model (Patterson et al., 2005; Quinn & Rohrbaugh, 1983). The Competing Values model consists of four quadrants representing four different managerial ideologies. According to the model these ideologies differ with regard to their criteria for efficiency along two dimensions. These

two dimensions are flexibility versus control, and internal versus external orientation. The four ideologies are associated with different outcomes. They also correspond to four major traditions in management. *The human relations* approach is characterised by having an internal focus and a flexible orientation towards the environment. This managerial approach emphasises well-being, growth and commitment. As such, it has a focus on the employees of the organization. *The internal process approach* to management also has an internal focus, but emphasises tight control over the employees. In so doing, it focuses on formalization and internal control. *The open systems approach* on the other hand, has an external focus and focuses on a flexible relationship with the environment. This managerial ideology stresses the importance of adapting the organization to the environment. Innovation is considered central to adaption. The fourth quadrant is represented by *the rational goal approach*. It is characterized by an external focus and tight control within the organization. This managerial approach emphasizes the attainment of organizational goal and focuses on productivity. By using the Competing Values model as a basis for selecting climate dimensions, the OCM is intended to have a more robust theoretical foundation. The dimensions are linked to managerial approaches that have long running traditions with work- and organizational psychology. By having different criteria for efficiency, these four approaches will lead to particular behaviour in employees. Organizations will in practice not adhere to a single quadrant, but emphasized elements of these managerial approaches to different degrees. Therefore it is likely that any given organization implements aspects from all of the four quadrants.

The 17 dimensions that constitute the OCM are distributed between the four quadrants. The *human relations quadrant* consists of 6 dimensions: *Autonomy, Integration, Involvement, Supervisory support, Training and Welfare*. The *internal process quadrant* consists of the dimensions Formalization and Tradition. The open systems quadrant consists of three dimensions: *Innovation-and-Flexibility, Outward focus and Reflexivity*. Lastly, the *Rational goal quadrant* consists of 6 dimensions: *Clarity of organizational goals, Efficiency, Effort, Performance feedback, Pressure to produce and Quality*. These dimensions will be further defined in the methods-part.

In addition to being theoretically grounded, the OCM has also been empirically validated (Patterson et al., 2005). Confirmatory factor analysis was used to assess the structure of the measure and its generalizability. Tests indicated that the OCM had discriminant validity. Meaning that it could be used to discriminate between organizations. It

also had consensual validity, indicating that employees within the same organization tended to assess the work environment in a similar fashion. For example, management and non-management tended to assess the environment in the same way. Concurrent validity was tested by comparing results from questionnaires tapping the 17 scales with semi-structured interviews conducted before the administration of the questionnaires. Predictive validity of the OCM was conducted a year after the administration of the questionnaire. This consisted of examining whether organizational innovation was associated with the ratings given by employees a year earlier. At a whole, this indicates that the OCM has been empirically validated on a range of areas that are important for its practical use within research.

Prior mixed-method studies.

Three known studies have compared quantitative questionnaires using close-ended questions and quantitative interviews using open-ended questions. Two of these have examined stress, while the last is a prior master thesis based on parts of the data being used in the current study.

The aim of a study carried out by Jex, Adams, Elacqua and Lux (1997) was to investigate whether questionnaires using close-ended questions were accurately measuring the actual experiences of occupational stress in employees. Their sample consisted of 151 females clerks. The participants completed quantitative questionnaires with items that tapped three common stressors. These stressors were *role ambiguity, role conflict and interpersonal conflict*. In the study the respondents were also measured on psychological and physical strain and on self-esteem. The participants experience of stressors was also measured using a qualitative method. This method consisted of asking the participants to recall a recent work related incident associated with each the three aforementioned stressors. They were instructed that this incidents should be illustrative of the degree they experienced the given stressor in their work environment. The incident should be described with as much details as the respondents themselves considered relevant. The qualitative and quantitative measures of stressors were compared by using a multitrait-multimethod analysis. While the results indicated a convergence between the two measures, it also revealed a possible bias on the qualitative measure due to the ordering of the items. The two measures were also compared by correlating them with the measure of self-esteem. The results indicated that only the qualitative measure was associated with self-esteem. Lastly, the researchers used regression analyses to see if the qualitative and quantitative measures explained equal amounts of variation in strain. Though both measures explained a considerable amount of variation, the

quantitative accounted for a greater proportion. The authors of the study concluded that qualitative measures could add to an understanding of stress by providing a more detailed picture of the phenomena of interest.

The second study that has compared close-ended questionnaire items and open-ended questions was carried out by Mazzola, Walker, Shockley and Spector (2011). This study used the same qualitative and quantitative types of measures as Jex et al (1997). One difference was that the respondents were asked to describe one incident of the three stressors of interest. In the study by Jex et al (1997) the respondents were to identify incidents for all stressors. The aim of this study was to see if the quantitative scales of stressors could discriminate between the participants who reported corresponding types of stressors on the qualitative measure and those that did not. The stressor being measured were *interpersonal conflict, constraints* and *quantitative workload*. The sample consisted of 207 graduate assistants. Results showed that the quantitative scales for *constraints* and *workload* could discriminate between participants who reported the corresponding stressor on the qualitative measure and those who reported no stressors at all. However, the quantitative scales could not discriminate between those who reported the stressor and those that reported a different stressor. The authors of this study also concluded that using a combination of qualitative and quantitative measures provided a more detailed picture of stressful events.

While both of these studies indicate that qualitative and quantitative measures can display converging results, they also show that the measures do diverge. Both converging and diverging results can be valuable (Lund, 2011). Converging results from alternative approaches attests to the validity of results. Diverging results on the other hand, can provide the researcher with new knowledge on the area being investigated and lead to new hypothesis.

The third known study that compares quantitative close-ended questionnaires with open-ended questions is a master thesis based on parts of the same sample being used in the current study. Berglund (2012) compared data gathered using the OCM questionnaire with data gathered from open-ended semi-structured interviews. The sample consisted of 33 participants employed in the Norwegian police organization. In this study data gathered from interviews were unitized and top-down coded on each of the scales of the OCM. The aim of the study was to see if the qualitative and quantitative measures could produce converging results. The two sets of data was compared by executing a series of paired-sample t-tests. As the means on 13 of the 17 scales were not significantly different, the study indicated that the two measures could provide converging results.

In the study by Berglund (2012), data gathered using interviews had to be transformed in order to make a comparison with the scales on the OCM possible. This involved combining the number of positive and negative statements given by the participants into a difference score. The use of differences scores has several drawbacks as it makes certain assumptions about the relationship between the statements and scores on the OCM scales (Jeffrey R. Edwards, 2001). One assumption is that the positive and negative statements making up the difference has an equal but opposite effect on the scale in question. Another assumption is that different numbers of statements are expected to be associated with the same score on the questionnaire as long as the difference between positive and negative statements is equal. Edwards and Parry (1993) have argued that the combined use of polynomial regression and response surface methodology could overcome the problems inherent in the use of difference scores.

Response surface methodology.

Response surface analysis is an emerging technique within work- and organizational psychology (Shanock, Baran, Gentry, Pattison, & Heggestad, 2010). The technique has traditionally been used within the industrial world (Myers & Anderson-Cook, 2009). It was introduced into the context of work- and organizational psychology in the beginning of the 1990s as a response to the use of difference scores in organizational research (Jeffrey R. Edwards, 1995).

The technique can be used to study how different combinations of independent variables are related to a dependent variable. As such, it offers a more nuanced explanation of this relationship than what can be achieved using difference scores. The relationship is estimated using a quadratic regression equation where the independent variable (Z) is regressed on the two independent variables (X and Y) and their second order terms (X^2, XY and Y^2). The quadratic equation is as follows:

$$Z = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 X^2 + \beta_4 XY + \beta_5 Y^2 + \varepsilon$$

Prior to calculating the second order terms, the first order terms (X and Y) are centred. The purpose of this is to reduce the potentially high level of multicollinearity that is common to regression equations using orthogonal terms. The independent variables can be centred in different ways. For example, by using the median or the mean. Edwards and Parry (1993) recommends centring them at the midpoint of the scale being used in order to make the interpretation of the results easier.

The results from the polynomial regression are presented visually by a three dimensional response surface. This surface is estimated on the basis of the unstandardized regression coefficients from the polynomial regression. In order to examine how different combinations of the independent variables are related to the dependent variable, the slope and curvature of two lines crossing this surface is calculated using the unstandardized regression coefficients. These lines are the line of congruence and the line of incongruence. These lines can be tested to see if they are significantly different from zero.

One of the drawbacks of using polynomial regression with response surfaces, is that the surfaces can be difficult to interpret. In the following paragraphs an explanation will be given on how to interpret the different combinations of slopes and curvatures along the two lines. These explanations will be related to a hypothetical response surface. Figure 1a presents the hypothetical response surface.

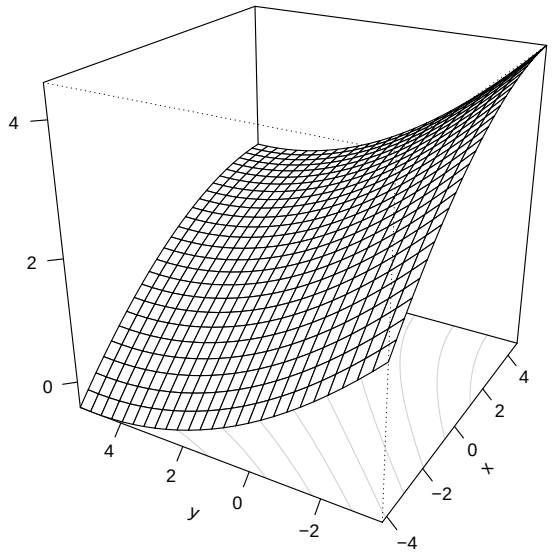

Figure 1a. Hypothetical response surface.

The line of congruence runs from the near most corner to the far most corner. In the near most corner both of the independent variables have a low value. In the far most corner both

independent variables have a high value. Moving along this line from the near most corner to the far most corner there is a gradual increase in value on both of the variables. If this gradual increase is associated with a gradual increase in the predicted values on the dependent variable, the surface will have a significant positive slope along the line of congruence. On the other hand, if this increase is associated with a gradual decrease in the predicted values on the dependent variable, the surface will have a significant negative slope along the line of congruence. A positive curvature along this line will give the surface a convex shape. Meaning that the predicted value on the dependent variable is expected to increase more sharply as one moves from the near most corner to the far most corner. A negative curvature will give the surface a concave shape along the line of congruence. This means that the surface will tend to flatten out as the line approaches the far corner. On the surface presented in figure 1 there is no increase in predicted values as one moves along the line of congruence. The predicted score seems to be 2 in each of the opposing corners. On the hypothetical surface there also seems to be no curvature along the line.

The line of incongruence runs perpendicularly to the line of congruence, from the leftmost corner to the rightmost corner of the plane. In the leftmost corner the independent variable Y has a high value and the independent variable X has a low value. In the rightmost corner the opposite is true. X has a high value, while Y has a low value. Moving along the line of incongruence from the left corner to the right corner, there is a gradual shift in the proportionality of values. From Y having an overweight in values in the left corner, to X having an overweight of values in the right corner. At the point where the line of congruence crosses the line of congruence both of the independent variables have the same value. They are congruent. This point, where the line of incongruence crosses the line of congruence is important, as it makes understanding the surface along the line of incongruence easier. Moving away from this intercept in either direction the level of incongruence increases. Moving away from the intercept in a leftward direction, the value of Y will increase compared to the value of X. Moving away from this intercept in a rightward direction the value of X will increase compared to the value of Y.

A significant curvature along the line of incongruence means that a discrepancy between the two independent variables is associated with a change in the dependent variable. A positive curvature means that the value of the dependent variable is predicted to increase as the degree of discrepancy between the two independent variables increase. Giving the surface a convex shape along the line of incongruence. Moving away from the point where the line of

incongruence crosses the line of congruence in a leftward or rightward direction, the values on the dependent variable are predicted to increase in either direction along the line of incongruence. And similarly, a negative curvature means that the dependent variable is predicted to decrease as the degree of discrepancy between the independent variables increase. Giving the surface a concave shape along the line of incongruence. A significant curvature in combination with a non-significant slope means that the predicted increase or decrease in the dependent variable is equal on both sides of the line of congruence. The line of incongruence will take on a parabolic shape.

If the line of incongruence has a significant positive slope in combination with a significant curvature, the value of the dependent variable is predicted to be higher when the discrepancy increases as one moves along the line of incongruence in a rightwards compared to a leftwards direction from the intercept with the line of congruence. A significant negative slope would indicate the opposite. That predicted values on the dependent variable would be higher following the line of incongruence in a leftwards direction compared to a rightward direction. As mentioned, using the intercept between the line of incongruence and the line of congruence as a reference point makes interpreting the surface easier. This is especially the case when the line of congruence has both a significant curvature and a significant slope.

When the line of incongruence has a significant slope without a significant curvature, it can be interpreted in the same way as the slope along line of congruence when it has no curvature. A significant positive slope indicates that the value on the dependent variable is predicted to increase as one moves along the line from the left corner to the right corner. And when the slope is significantly negative, the value on the dependent variable is expected to decrease as one follows the line from the left corner to the right corner. The relationship between the two independent variables and the dependent variable is linear.

On the surface presented in figure 1a there is no curvature along the line of incongruence. But the surface has a significant slope along the line. Moving along the line of incongruence from the leftmost corner to the rightmost corner the predicted score on the dependent variable is predicted to increase.

The Present Study

The aim of this study is to investigate if polynomial regression and response surface methodology can be used to study the relationship between statements from interviews and scores from questionnaire. In this study the number of positive and negative statements given by participants will represent the independent variables. This poses some challenges. Most previous studies using this technique have used independent variables that have been measured on scales that have the same maximum value (e.g. 5-point, 7-point scales). When using number of statements the maximum value can differ between the two independent variables. In addition to this, the first order variables (X and Y) has been centred at their mean value. As mentioned, Edwards and Parry (1993) recommends centring these variables at the midpoint of the scale the variables are measured on. While this makes it easier to interpret surfaces when using measures such as Likert-scales, it is problematic to do with statements. As the number of maximum statements on the independent variables will differ between the different categories being investigated, it would make it difficult to compare surfaces with each other. For this reason the independent variables have been centred at the mean of each variable. When the slope and curvature of the line of congruence and incongruence is tested, this is done at the point where both X and Y has the value of zero. By centring at the mean the slope and curvatures will be tested at the point where the average number of statements is located on the surface. Moving along the line in either direction from this point, we find combination of statements that are either over or under this average.

This can make it more difficult to interpret the surfaces. In order to ease interpretation a separate two dimensional contour plot will be included along with the three dimensional response surface. The use of contour plots is not common when using this technique in psychology. But some researchers have recommended using them (Cohen, Nahum-Shani, & Doveh, 2010), as they provide an additional tool to interpreting the surfaces. Figure 1b presents a contour plot of the hypothetical surface presented earlier. The contour plot can be studied the same way as one would study the three dimensional response surface. The curved lines represent the predicted values on the dependent variable. The solid line running from (close to) the bottom left corner to the top right corner represents the line of congruence. The dotted line running from the top left corner to the bottom right corner represents the line of incongruence. Following the line of congruence from the bottom left corner we see that the values on the dependent variable stay the same along the whole line. Following the line of

incongruence from the top left corner to the bottom right corner, we see the value on the dependent variable change as we move along the line.

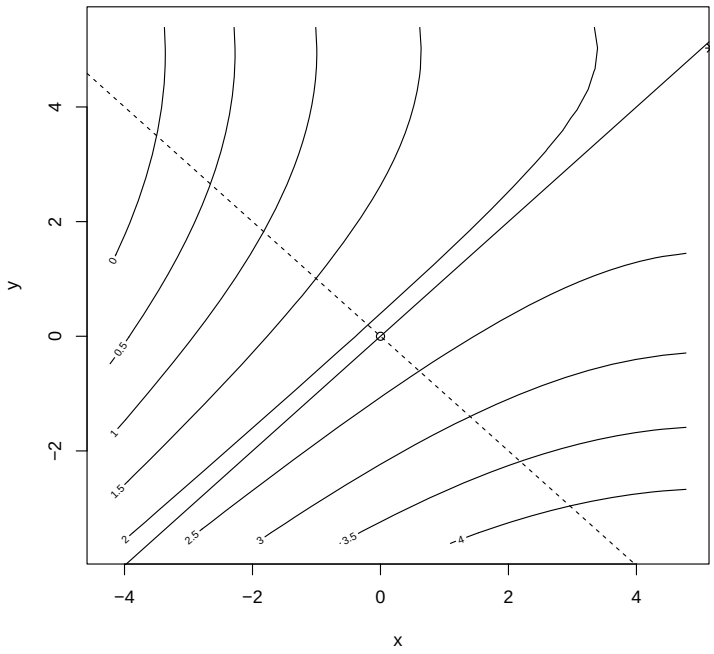

Figure 1b. Hypothetical Contour plot.

Aim of the study.

The aim of this study is to investigate if polynomial regression and response surface methodology can be used to study the relationship between statements from interviews and scores from questionnaires. In this study the number of positive and negative statements given by participants will represent the independent variables and the scores on the OCM will represent the dependent variable. The study uses a concurrent mixed model design where both the qualitative and quantitative strands of data are equally emphasized.

The research question of this study is:

Can polynomial regression and response surface methodology be used to study the relationship between statements from interviews and scores from questionnaires?

Method

Sample.

Qualitative sample. The Norwegian police organization consists of 27 districts. The qualitative sample is made up of 90 male and female participants strategically selected from all of the 27 districts. A strategic sampling procedure was chosen to ensure subsets corresponding to three different job levels within each of the districts. A letter was sent to the police commissioner in each of the districts, asking him or her to participate in the project. The commissioners were asked to select and arrange interviews with individuals corresponding to each the job levels of interest. Three interviews were conducted in each of the respective districts. There were some deviation from this plan. In four of the districts four interviews were conducted. In one of the districts eight interviews were conducted. The reason being that this district represented the largest district in the police organization. A decision was made that three interviews would not cover the nuances and differences between the different units in this district.

Quantitative sample. The quantitative sample emanated from the qualitative sample, as each interviewee was asked to fill out a survey after having been interviewed. The sample consists of 76 male and female participants. All but one of these were also a part of the qualitative sample (e.g. had been interviewed). All of the 27 districts were represented. The 51 first participants in the study received an email following the interview inviting them to complete the survey. 39 (76,5 %) of the 51 participants responded. In addition, we received a completed survey from a person who had not been interviewed. The last 38 participants were given a physical copy of the OCM survey after the completion of the interview. These participants were given the opportunity to fill out the survey and hand it in immediately or to do so later and send the completed survey by post using a stamped envelope accompanying the survey form. 36 of the 39 participants responded (92,3 %). At a whole, 75 of the 90 participants completed the survey, resulting in a response rate of 83,3 %.

Mixed sample. The mixed sample consists of 75 participants. The prerequisite for being included in this sample was that the participant had been interviewed and also returned a completed survey. All of the 27 districts were represented.

Measures.

Qualitative measures (open-ended interviews). Semi-structured interviews were conducted in accordance with the PEACE-model (Milne & Powell, 2010). Each interview consisted of four semi-structured open-ended questions founded on the SWOT format, whereby the interviewee was asked by the interviewer to reflect on the strengths, weaknesses, opportunities and threats relating to the quality of investigation in the Norwegian police. The purpose of the SWOT format was to provide the interviewee with a structure that facilitated reflection on the subject matter without forcing particular answers. This was assumed to be accomplished by encouraging free association within a multidimensional space set by the dimensions positive-negative and present-future. Follow-up questions were asked after the four questions. The follow-up questions encouraged the interviewee to clarify or elaborate on topics he or she had introduced during the course of the interview.

Quantitative measures (close-ended questionnaires). Quantitative data was gathered using the organizational climate measure (OCM). As previously mentioned, the OCM consists of 82 items that assesses respondents perception of their work environment on 17 scales. These scales are theoretically grounded in the 4 managerial ideologies of the Competing Values model. This study used a version of the questionnaire that had been translated into Norwegian. The questionnaire was modified to accommodate the Norwegian police organization. The modification consisted of changing several words assumed to be incomprehensible within the given context.

The 17 climate scales of the OCM are: 1. *Autonomy*: designing job in way which give employees wide scope to enact work. 2. *Integration*: the extent of interdepartmental trust and cooperation. 3. *Involvement*: employees have considerable influence over decision-making. 4: *Supervisory support*: the extent to which employees experience support and understanding from their immediate supervisor. 5: *Training*: a concern with developing employee skills. 6: *Welfare*: the extent to which the organization values and cares for employees. 7: *Formalization*: a concern with formal rules and procedures. 8: *Tradition*: the extent to which established ways of doing things are valued. 9: *Innovation & flexibility*: flexibility - an orientation toward change, innovation – the extent of encouragement for new ideas and innovative approaches. 10. *Outward focus*: the extent to which the organization is responsive to the needs of the customer and the marketplace in general. 11. *Reflexivity*: a concern with reviewing and reflecting upon objectives, strategies and work processes, in order to adapt to the environment. 12. *Clarity of organizational goals*: a concern with clearly defining the goals

of the organization. 13. *Efficiency*: the degree of importance placed on employee efficiency and productivity at work. 14. *Effort*: how hard people in the organization work towards achieving goals. 15. *Performance feedback*: the measurement and feedback of job performance. 16. *Pressure to produce*: the extent of pressure for employees to meet targets. 17. *Quality*: the emphasis given to quality procedures.

Procedure.

Interviews. Interviews were conducted by an experienced interviewer with extensive knowledge of the police organization. All interviews were conducted by the same interviewer to prevent variation in responses across interview situations. This interviewer was on all but a few occasions accompanied by a second interviewer. Interviews were conducted in accordance with an interview guide. The interview would start with the interviewee explaining his or her background. Next the interviewee was encouraged to reflect on the quality of police investigation by being asked the four questions based on the SWOT format. The four questions were asked in the same successive order for each interview, but the interviewee was free to deviate from this order at his or her own initiative. In the event of any follow-up questions, these were asked after the interviewee had indicated that he or she had nothing more to say. Prior to concluding the interview, the interviewer would give the interviewee a last opportunity to add any additional information should he or she wish to. The interview would be concluded according to the PEACE model, giving the interviewee an opportunity to express his or her thoughts about being interviewed.

Surveys. For the first 51 participants in the study, the survey was sent by mail after they had been interviewed. The letters were sent to the participants at their place of work in personally addressed envelopes. The letters contained a personal letter, an instruction, the questionnaire and an addressed return envelope. The letters were sent out by the police postal system on 2. December 2011. The collection of data ended on the 2. February 2011. The last 38 participants were given the questionnaire directly following the interview. They were given an opportunity the fill out the questionnaire and hand it in to the interviewer immediately, or to send it in later using an addressed return envelope.

Transcription. Recordings from the interview were transcribed verbatim in Norwegian Bokmål, including words like “ehm” and “mhm”. This meant that dialect was translated, and that potentially some connotations related to the respective dialect was lost. The majority of the first 51 interviews were transcribed by research assistants, with the remaining 12 interviews being transcribed by master students involved in the research project

during the fall of 2011 and spring of 2012. The 38 interviews from the current phase of the project, were distributed between and transcribed by the author and 6 fellow students. In order to achieve a shared understanding of the transcription process, all seven student transcribed three of the interviews. After the transcription the students screened the transcriptions for errors while listening to the recording. The screening involved detecting and counting the number of spelling mistakes and omissions (e.g. missing words). En evaluation of the process concluded with that the omissions detected had no bearing on the information being communicated by the interviewee. The differences in the transcriptions primarily consisted of number of detected non-semantic units as “mhm” and “ehm”.

Content analysis. In order to make the qualitative data from the transcripts comparable to the quantitative data from the questionnaires, the transcripts were unitized and top-down coded according to the principles of content analysis (Krippendorff, 2004; Neuendorf, 2002).

Preparations and coding scheme. Prior to performing the content analysis, the author and fellow students carried out an extensive training process. During this process all the students performed all the steps in the content analysis on three different interviews. The result of the process was a formal coding scheme. This coding scheme outlined how the different steps in the content analysis should be performed in the subsequent analysis performed by the individual students. For the coding scheme see appendix C.

Unitizing. The process of unitizing involves dividing the transcript for each of the interviews into a multitude of distinct units. 52 interviews had been unitized at a previous stage of the project. The 38 new interviews were unitized by the author and fellow students. The process of unitizing involves dividing the transcript for each of the interviews into a multitude of distinct units. A total number of 19 052 statements were identified in the transcripts from the 90 interviews. Reliability among the seven students on unitizing was measured during the preparation stage. All students unitized three different transcripts in consecutive order. Apart from the first transcript, there was little discrepancy between the students with regard to number of statements extracted from the transcripts.

Top-down coding. Each of the statements in the qualitative sample was coded on three models: SWOT, IGLO and OCM. The coding on SWOT and IGLO was done concurrently. The coding on OCM was done after the conclusion of the SWOT and IGLO coding. During the preceding phase of the project, the majority of the coding of the statements on SWOT and IGLO had been conducted by research assistants, and statements

corresponding to the 51 interviews had been coded on OCM by two master students. The 38 interviews conducted in the current phase of the project was coded on SWOT and IGLO by the author and a fellow student. The statements corresponding to the new 38 interviews were coded on the different levels of the OCM-measure by the author.

SWOT. Each statement in the qualitative sample was coded on one of the 4 SWOT categories. These were: 1. *Strengths*: positive statements concerning the "here and now". 2: *Weaknesses*: negative statements about the "here and now". 3. *Opportunities*: positive statements regarding the future. 4. *Threats*: negative statements regarding the future. Statements that did not fit any of these categories were coded as residuals. Coding was done by first identifying whether the statement was positive or negative, before deciding whether the statement was directed toward the presence or the future.

IGLO. Each statement in the qualitative sample was also coded on the level of the organization the statement was directed at. These were: 1. *the individual level*: individual perceptions, feelings and opinions, 2. *the group level*: interaction and cooperation in work groups, teams and departments, 3. *the leadership level*: behaviour of immediate supervisors, other leaders or top management and 4. *the organizational level*: management practices, organizational culture, strategies, organizational goals and values, and the physical environment of the organization. Statements that did not fit any of these categories were coded as residuals.

OCM. Each statement in the qualitative sample was coded on one of the 17 categories of the OCM. Statements that did not fit any of the categories were coded as residuals. The hypothetical statement "the people in my group enjoy working together" would have been coded as a residual, since no one of the OCM factors cover intradepartmental cooperation. Coding was done by first attempting to identify the quadrant the statement belonged to. The purpose of doing this is to reduce the number of potential categories from 17 to a more manageable number. Doing so can also facilitate more consistent coding on categories that otherwise might seem to overlap. An example is whether a statement should be coded on "Innovation" or "Tradition". In cases such as this, it might be easier to identify the correct quadrant, than to decide which category best fits the statement. How the statement had previously been coded on SWOT and IGLO was also taken into consideration. This is important, since a statement that can be seen as positive during top-down coding on OCM, not necessarily was seen as positive during the coding on SWOT and IGLO. Coding on OCM can make the coder focus on different parts of the statement. For a description of the 17 OCM

categories see the previous section ("measures"). For the coding scheme used, see appendix D. The items from the OCM questionnaire was used in the beginning of the process to understand the content of each category, but they did not play a pivotal role during the actual coding.

Data treatment and statistical analysis.

Interview. All statements were transferred to the statistical software SPSS 20. The statements were entered in "long" or "molten" form, meaning that each statement represented a distinct case and was represented by a single row. Separate columns represented the level each statement was coded on with regard to the SWOT, IGLO and OCM-models. In addition to these, separate columns also contained the identification number given to interviewee who the statement belonged to, along with columns containing the corresponding identification number given to the district and job level. The resulting data file contained no information that would make it possible to link any of the information, district or job level to the interviewee or the district he or she worked in, or the job level.

The mixed strand of data was aggregated at the level of each participant. This involves a summing up of the number of statements coded as strengths and weaknesses on SWOT, on each of the top-down coded categories on OCM, for each participant. Hence, in the resulting aggregated file, each case (row) would represent a single participant, and separate variables (columns) would represent the sum of strengths and weaknesses for each participant on each of the 17 OCM –categories (34 variables in total). An additional 17 variables (columns) were included in the file, each representing the score each participant had on each of the 17 factors on the OCM questionnaire.

For example, a hypothetical participant in the study could have talked about autonomy for 5 minutes during the interview. These 5 minutes could then have been separated into 10 statements during the unitizing process. During the top-down coding on SWOT, 7 of these statements could have been coded as strengths and 3 of the statements been coded as weaknesses. During the top-down coding on OCM, all 10 statements were correctly identified and coded as "Autonomy". The participant also completed the OCM questionnaire and got an aggregated scored of 4 on the OCM category "Autonomy". In the final data file, this participant would represent a single case (row) with the value "7" on the variable "Strengths Autonomy", the value "3" on the variable "Weaknesses Autonomy" and the value "4" on the variable representing the score on the category "Autonomy" on the OCM questionnaire.

Survey. The answers from the questionnaire were manually transferred into the

statistical software SPSS 20. Missing values were replaced by the mean on the respective scale. The values on the respective scales for each respondent were combined in order to generate an aggregated score on each scale for each respondent. This was done in accordance with the guidelines provided by Patterson et al (2005).

Inter-rater reliability. Inter-rater reliability among seven coders was measured by calculating pairwise percentage agreement and Krippendorff's Alpha across three different interviews. Reliability was measured both for top-down coding on SWOT and top-down coding on IGLO and top-down coding on OCM.

On SWOT pairwise percentage agreement ranged from 57,84% to 73,69%. With an average percentage agreement of 67,12%. Krippendorff's Alpha ranged from 0,44 to 0,66, with an average alpha of 0,57.

On IGLO pairwise percentage agreement ranged from 48,94% to 57,55%, with an average agreement of 54,11%. Krippendorff's Alpha ranged from 0,26 to 0,40, with an average alpha of 0,35.

On OCM inter-rater reliability between the author and the students who coded on OCM during the preceding stage of the project was based on a single interview. Inter-rater was measured in the middle of the coding process. Pairwise percentage agreement ranged from 45,13% to 52,21% between the three coders, with an average pairwise percentage agreement of 49,26%. Krippendorff's Alpha was 0,36. This represents only a marginal difference from the reliability estimates from the previous stage of the project. The students who coded on the OCM in the preceding stage reported an average Krippendorff's Alpha of 0,41, with an average pairwise percentage agreement of 50.83%. Pairwise percentage agreement and Krippendorff's Alpha was calculated using the dfreelon ReCal online reliability calculator (Freelon, 2010).

Analysis.

Response surface methodology. Polynomial regression and response surface methodology was used to study how the relationship between positive and negative statements from interviews was related to thematically related scales on the OCM measure. The response surfaces represent the point of integration in the mixed method design, where the qualitative and quantitative strands of data are integrated.

Robust regression analysis. The response surfaces were estimated using the unstandardized coefficients from polynomial regression analyses. In total, a series of 17 regression analyses were conducted. Due to the presence of outliers a series of robust

regression analyses were performed. The robust regression analyses used an MM-estimator using a bi-square redescending score function (Yohai, 1987). It used an initial S-estimator (Rousseeuw & Yohai, 1984) which was computed using the Fast-S algorithm of Saliban-Barrera and Yohai (Salibian-Barrera & Yohai, 2006). This type of robust regression analysis was chosen due to its robustness when dealing with multiple outliers (Wilcox, 2012).

Regression analysis was used on each of the 17 subsets of the mixed data sample. Each subset consisted of participants who had a minimum of one positive or negative statement that was coded on the OCM category corresponding to the OCM factor the interview data was being related to. Meaning that, if we want to study how positive and negative statements from interviews on the topic of autonomy are related to the scale Autonomy on the survey, we only select statements from the interview that have been top-down coded on the category Autonomy during the coding process. Each of the 17 subsets only contained statements from the interview related to one of the 17 factors on the OCM questionnaire.

Screening data for outliers and influential cases. Regression and response surface estimates are sensitive to outliers and influential cases (Jeffrey R Edwards & Rothbard, 1999). In smaller samples a few observations can have a large impact on these estimates. This is can also be the case in larger samples (Bollen & Jackman, 1990). In the current study the regression equations were screened using studentized residuals, adjusted predicted value DFFit (Belsey, Kuh, & Welsch, 1980), Cook's distance (Cook, 1977) and leverage (hat-values). When investigating Cook's distances, the cut-off values where calculated using the formula:

$$\frac{4}{n}$$

Where *n* is the number of observations. This cut-off value has been suggested by Bollen and Jackman (1990). Observations with a Cook's D greater than the cut-off value were considered potentially problematic. The cut-off value for leverages was calculated using the formula:

$$2\left(\frac{p}{n}\right)$$

Where *p* is the number of estimates (number of predictors + intercept), and *n* is the number of observations. This cut-off value is in accordance with the recommendation by Hoaglin and

Welsch (1978). Hat values greater than the cut-off value was considered to exert potentially excessive influence.

Controlling Type I error. Following Edwards and Rothbard (1999), the Holm–Bonferroni method ("sequentially rejective Bonferroni test") was used to control for familywise error rate (Holm, 1979). The method was implemented by sorting the probability estimates from the F-tests and Wald-tests in ascending order and multiplying the first value with 17, the second value with 16, and so forth until one of the values exceeded the necessary alpha level of 0,05.

Software. SPSS 20 was used to aggregate data and for descriptive statistics. OLS-regression and robust regression was done using R version 2.14.2. Robust regression was done using the package "robustbase".

Ethical considerations. Participation in the study was based on informed consent. Prior to the interviews all participants received a letter outlining the procedure and informing them of their rights of confidentiality and right to withdraw from the study. The qualitative and quantitative strands of data were kept separate and the key to linking the datasets was restricted. Due to the nature of the data, NSD (Norsk samfunnsvitenskapelig datatjeneste) was not affiliated.

Results

Descriptive statistics.

Qualitative sample. The qualitative data sample consisted of 19052 statements extracted from interviews with 90 participants that had been top-down coded on SWOT and OCM. Table 1 shows the distribution of statements coded on SWOT distributed on the 17 categories of the OCM.

Table 1

SWOT-statements distributed on OCM

	SWOT					
OCM scale	Strengths	Weaknesses	Opportunities	Threats	Residual	Total
1. Autonomy	81	49	30	9	6	175
2. Integration	856	530	417	96	32	1931
3. Involvement	137	92	102	42	9	382
4. Support	131	85	75	15	4	310
5. Training	480	388	560	141	17	1586
6. Welfare	106	174	125	118	8	531
7. Formalization	270	214	137	82	20	723
8. Tradition	67	203	46	93	7	416
9. Innovation	170	114	246	52	7	589
10. Focus	150	81	98	88	7	424
11. Reflexivity	207	181	357	87	7	839
12. Clarity	131	99	113	94	5	442
13. Efficiency	305	340	265	105	11	1026
14. Effort	230	105	69	31	5	440
15. Feedback	217	197	183	50	12	659
16. Pressure	44	229	50	120	5	448
17. Quality	207	252	217	98	5	779
18. Residual	1886	1669	1046	809	1942	7352
Total	5675	5002	4136	2130	2109	19052

A total of 16943 statements (88,9 %) could be coded as strengths, weaknesses, opportunities or threats, while 2109 statements (11,1 %) were coded as residuals. This indicates the OCM categories were capable of tapping a large proportion of the statements given by participants during interviews. The majority of the statements were directed at the present, with a total of 10677 statements (56,0 %), while 6266 statements (32,9 %) were directed at the future. The highest frequency of statements could be found on the category Strengths, with a total of 5675 statements (29,8 %). Statements coded as Weaknesses followed shortly, with a total of 5002 statements (26,2 %). The third most frequent category of statements on SWOT was Opportunities, with a total of 4136 statements (21,7 %). This category had almost twice the frequency of the least frequent category, Threats, which had a total of 2130 statements (11,2 %).

A total of 11700 statements (61,4 %) could be coded on one of the 17 categories of the OCM. Statements that did not fit any of these categories were coded as residuals. The residuals consisted of 7352 statements (38,6 %). During the preceding stage of the project, statements coded as residuals on SWOT were not coded on the OCM. This amounted to a total number of 1358 statements. In the current sample these statements were recoded as residuals on OCM.

Comparing the proportion of statements coded as residuals across the four SWOT-categories, show that an is little difference between positive and negative statements directed at the present, and a higher difference between positive and negative statements directed at the future. A total of 1886 (33,2 %) coded as Strengths were coded as residuals on OCM, while 1669 statements (33,4 %) coded as Weaknesses were coded as residuals. This gives a difference of only 0,2%. A total of 1046 statements (25,3 %) of statements coded as Opportunities were coded on OCM, compared to 2130 statements (38,0 %) coded as Threats. This amounts to a difference of 12,3 % of the statements.

The OCM-category with the greatest number of statements across all categories of the SWOT is Integration, with a total of 1931 statements (10,1 %). On this category we also found the highest frequency of statements coded as Strengths and Weaknesses on SWOT, with 856 and 530 statements respectively. The OCM-category with the second largest number of statements is Training, with 1586 statements (8,3 %). Apart from having the second largest concentration of statements coded as Strengths and Weaknesses, this category also contains the greatest number of statements coded as Opportunities and Threats, with a total of 560 and 141 statements. The OCM-category with the least number of statements is Autonomy with

only 175 statements (0,9 %). This is also the category where we found the lowest concentration of both Weaknesses, Opportunities and Threats, with 49, 30 and 9 statements. The smallest number of statements coded as Strengths was found on the OCM-category Pressure, with only 44 statements compared to 229 statements coded as Weaknesses.

The first six categories on the OCM constitute the human relations quadrant. This was the quadrant with the greatest number of statements, with a total of 4915 statements (25,8 %). Within this quadrant we also found the two largest OCM-categories with regard to number of statements. Integration with 1931 statements and Training with 1586 statements. The category Autonomy also belongs to this quadrant, and represents as previously mentioned the smallest OCM-category. The second largest quadrant is Rational Goal, consisting of a total of 3794 statements (19,9 %) spanning the OCM-categories 12 through 17. The categories Innovation, Focus and Reflexivity belong to the quadrant representing the open systems approach. This quadrant consisted of 1853 number of statements (9,7 %). The quadrant with the least number of statements is the internal process quadrant, with 1139 statements (6,0 %).

Quantitative sample. The quantitative sample consisted of data extracted from the OCM questionnaires received from 76 participants. Table 2 presents the descriptive statistics for the quantitative sample.

The OCM questionnaire consists of 82 items tapping the respondents assessment of 17 scales. The questionnaire uses a 4 point Likert scale as response format. The 4 points on the Likert scale was 1: "*definitely false*", 2:"*mostly false*", 3: "*mostly true*" and 4: "*definitely true*". An example of an item aimed at tapping the scale (Supervisory) Support is: "Supervisors here are really good at understanding peoples' problems". Units tapping the same scale were aggregated to give each participant an aggregated score on each of the scales. An aggregated score below 2 is considered negative, a score above 3 is considered positive, and a score between 2 and 3 is considered neutral. The average score on 14 of the scales were within the neutral range spanning scores between 2 and 3. None of the scales had a mean score below 2. On three of the scales the mean was slightly above 3. These three dimensions were Welfare (M=3,12), Formalization (M=3,15) and Effort (M=3,05). The dimension Efficiency (M=2,02) was on average rated lowest. The standard deviations were similar across scales, with a range from 0,31 to 0,66. At a whole, the quantitative data gathered indicated that the participants on average were neutral towards the aspects of the work climate measured by the questionnaire and none of the scales stood out as being particularly negative.

Table 2

Descriptive statistics for the quantitative sample

OCM Scale	N	Min	Max	M	SD
1. Autonomy	76	1,60	3,60	2,91	0,33
2. Integration	76	1,40	4,00	2,83	0,51
3. Involvement	76	1,80	3,67	2,91	0,46
4. Support	76	2,00	3,80	2,94	0,31
5. Training	76	1,25	3,75	2,53	0,54
6. Welfare	76	1,25	4,00	3,12	0,54
7. Formalization	76	2,00	4,00	3,15	0,51
8. Tradition	76	1,00	4,00	2,62	0,66
9. Innovation	76	1,00	3,67	2,36	0,49
10. Focus	76	1,60	4,00	2,82	0,54
11. Reflexivity	76	1,60	3,40	2,62	0,40
12. Clarity	76	1,40	3,60	2,61	0,51
13. Efficiency	76	1,00	3,50	2,02	0,50
14. Effort	76	1,60	4,00	3,05	0,42
15. Feedback	76	1,20	3,60	2,52	0,52
16. Pressure	76	1,40	3,80	2,48	0,47
17. Quality	76	1,50	4,00	2,93	0,47

The OCM questionnaire consists of 82 items tapping the respondents assessment of 17 scales. The questionnaire uses a 4 point Likert scale as response format. The 4 points on the Likert scale was 1: "*definitely false*", 2:"*mostly false*", 3: "*mostly true*" and 4: "*definitely true*". An example of an item aimed at tapping the scale (Supervisory) Support is: "Supervisors here are really good at understanding peoples' problems". Units tapping the same scale were aggregated to give each participant an aggregated score on each of the scales. An aggregated score below 2 is considered negative, a score above 3 is considered positive, and a score between 2 and 3 is considered neutral. The average score on 14 of the scales were within the neutral range spanning scores between 2 and 3. None of the scales had a mean score below 2. On three of the scales the mean was slightly above 3. These three dimensions were Welfare (M=3,12), Formalization (M=3,15) and Effort (M=3,05). The dimension

Efficiency (M=2,02) was on average rated lowest. The standard deviations were similar across scales, with a range from 0,31 to 0,66. At a whole, the quantitative data gathered indicated that the participants on average were neutral towards the aspects of the work climate measured by the questionnaire and none of the scales stood out as being particularly negative.

Mixed data sample. The mixed sample consisted of 8617 statements distributed on 75 participants. A total of 5749 statements could be coded on one of the 17 OCM categories (66,7 %). A total of 2868 statements (33,3 %) had been coded as residuals. The mixed data sample was extracted from the qualitative sample by extracting only statements that had been coded as Strengths or Weaknesses on SWOT on participants that had also returned a completed OCM questionnaire. From this sample 17 subsets were extracted. Each corresponding to one of the 17 scales on the OCM. For each sample only statements that had been coded on the OCM-category corresponding to the respective OCM scale was extracted. These statements were aggregated on an individual level. This means that the subsample for Autonomy consisted of 36 cases, each of which corresponded to a specific participant who had at least one statement coded as a Strength or a Weakness on the category Autonomy. Each participant was paired with his or her corresponding aggregated score on the OCM questionnaire for the respective scale. Table 3 presents the descriptive statistics for each of the 17 subsets.

The subset with the greatest number of statements was the Integration sample, with a total of 1107 statements distributed on 75 participants. The average number of statements coded as Strengths per participant was 8,80 statements, and the average number of statements coded as Weaknesses was 6,52. The subset with the smallest number of statements was Autonomy, with a total of 101 statements. These statements belonged to 36 participants, who on an average was coded on 1,69 statements on Strengths and 1,11 statements on Weaknesses. Compared to the Integration subsample, the average score on the OCM scale from the questionnaire was a little higher, with an average score of 2,90 on the Autonomy scale and 2,84 on the Integration scale. In 10 of the 17 subsets the average number of positive statements was greater than the number of negative statements. A large proportion of the subsets where negative statements on average outnumbered positive were within the Rational goal quadrant. In both the subset for Clarity of organizational goals, Efficiency, Pressure to produce and Quality procedures there was on average more statements coded as Weaknesses than Strengths. This amounts to 4 out of 6 OCM subsets. In comparison, only the subset with statements coded on Welfare had an overweight of negative statements in the Human relations

quadrant. This is notable, since Welfare was one of only three dimensions that on average was rated as positive on the OCM questionnaire. All of the three subsets in the Open systems quadrant had an overweight of positive statements. In all of the subsets the standard deviation on both Strengths and Weaknesses was high. This indicates that there was on average a large variation between participants with regard to number of positive and negative statements. The means and standard deviations on the scores from the OCM questionnaire were similar to the ones observed in the quantitative sample.

Table 3

Descriptive statistics for the 17 subsets extracted from the mixed sample

		Interview				Questionnaire	
		Strengths		Weaknesses		OCM scores	
OCM Scale	N	M	SD	M	SD	M	SD
1. Autonomy	36	1,69	1,77	1,11	1,65	2,90	0,33
2. Integration	75	8,80	6,52	5,96	4,86	2,84	0,51
3. Involvement	57	1,75	3,04	1,46	1,63	2,93	0,47
4. Support	43	2,28	2,55	1,63	1,85	2,99	0,29
5. Training	74	5,08	5,29	4,14	3,33	2,53	0,54
6. Welfare	59	1,29	1,65	2,46	2,42	3,07	0,56
7. Formalization	70	2,84	3,86	2,63	2,59	3,15	0,52
8. Tradition	36	1,47	2,95	4,61	6,06	2,39	0,65
9. Innovation	53	2,26	2,92	1,91	2,14	2,42	0,49
10. Focus	48	2,27	2,33	1,10	1,04	2,81	0,54
11. Reflexivity	56	2,91	3,80	2,68	2,94	2,60	0,41
12. Clarity	50	1,66	1,92	1,82	1,87	2,61	0,52
13. Efficiency	73	3,56	2,80	4,05	3,63	2,03	0,50
14. Effort	68	3,12	3,35	1,35	2,06	3,05	0,40
15. Feedback	59	3,05	4,26	2,95	2,45	2,43	0,49
16. Pressure	58	0,55	0,94	2,93	2,32	2,53	0,45
17. Quality	67	2,64	3,31	3,31	3,34	2,94	0,49

Inferential statistics.

A series of 17 least square regression analyses were conducted. Table 4 shows the results from the regression analyses.

Table 4

Results from the OLS-regression analyses

		F-test [a]	Unstandardized coefficients				
OCM scale	N	F	X	Y	X^2	XY	Y^2
1. Autonomy	36	1,70	0,11	0,01	-0,03	0,03	0,02
2. Integration	75	0,07	0,00	-0,01	0,00	0,00	0,00
3. Involvement	57	0,80	0,08	-0,04	-0,01	0,03	0,01
4. Support	43	0,17	0,01	-0,01	-0,01	0,00	0,00
5. Training	74	2,16	0,02	-0,03	0,00	0,00	-0,01
6. Welfare	59	2,63	0,18**	-0,03	-0,03	0,03	0,02
7. Formalization	70	1,57	-0,02	-0,07	0,00	0,01	0,01
8. Tradition	36	0,50	-0,09	0,01	0,01	0,00	0,00
9. Innovation	53	0,90	0,03	0,02	0,00	-0,03	0,00
10. Focus	48	3,22	-0,03	-0,27**	0,01	-0,03	-0,05
11. Reflexivity	56	0,73	0,06	-0,02	0,00	-0,01	0,00
12. Clarity	50	1,22	0,16*	0,04	-0,06*	0,02	-0,02
13. Efficiency	73	0,91	-0,04	-0,01	0,00	0,01	0,00
14. Effort	68	1,06	-0,03	0,06	0,00	0,01	-0,02
15. Feedback	59	4,31*	0,05	-0,11***	0,00	0,00	0,02**
16. Pressure	58	0,76	-0,10	0,01	0,04	-0,01	0,01
17. Quality	67	1,26	0,06	0,01	0,00	0,00	0,00

Note. [a] Probability estimates were corrected for family-wise error using the Holm-Bonferroni method.
* p< .05. ** p< .01. *** p< .001.

The results showed that the polynomial model could only accounts for a significant amount of variation on 1 of the 17 scales. This scale being Feedback, $F(5,53) = 4,31$, $p < 0,1$. Since outliers and influential cases can have a large impact on the results of polynomial regression models, these were tested for using studentized residuals, adjusted predicted value (DFFit), Cook's distance and leverage (hat-values). Table 5 presents the maximum and minimum values of the studentized residuals and adjusted predicted values. The table also shows the

maximum value of Cooks distance, along with the cut-off value and number of cases in each subset that exceeded the cut-off value.

Table 5

Results from the test of outliers

	Studentized residuals		Cooks distances			DFFits	
Scale	Max	Min	N [b]	Cut-off [a]	Max	Max	Min
1. Autonomy	1,23	-6,66	2	0,13	0,25	1,23	-1,72
2. Integration	1,69	-2,84	7	0,06	0,48	1,69	-0,87
3. Involvement	1,44	-2,72	3	0,08	53,89	1,44	-18,06
4. Support	1,01	-4,16	4	0,11	0,71	1,01	-2,08
5. Training	0,97	-2,91	6	0,06	0,30	0,97	-1,35
6. Welfare	2,85	-3,50	6	0,08	1,33	2,85	-1,64
7. Formalization	0,56	-2,73	1	0,06	3,20	0,56	-4,41
8. Tradition	1,72	-1,94	6	0,13	16,56	1,72	-10,25
9. Innovation	4,10	-2,77	3	0,09	2,82	4,10	-0,91
10. Focus	5,92	-2,39	3	0,10	5,59	5,92	-0,78
11. Reflexivity	3,69	-2,69	4	0,08	5,47	3,69	-5,89
12. Clarity	2,58	-1,87	3	0,09	1,11	2,58	-1,08
13. Efficiency	0,94	-2,28	3	0,06	0,15	0,94	-0,82
14. Effort	0,56	-3,84	2	0,06	0,53	0,56	-1,77
15. Feedback	1,54	-3,23	5	0,08	1,62	1,54	-3,11
16. Pressure	0,90	-2,23	5	0,08	0,79	0,90	-2,17
17. Quality	0,53	-2,81	2	0,07	0,29	0,53	-1,31

Note. [a] Cut-off value calculated using the formula 4/n. [b] Number of cases exceeding cut-off value.

The test indicated that all of the scales had cases with potentially outlying values. As there was no reason to assume these cases were a result of measurement error or that the data could have been entered wrongly into the data files, a series of robust regression analyses were conducted on the 17 subsets. Table 6 shows the results from the robust regression analyses.

Table 6

Results from the robust regression analyses

OCM scale	N	Wald-test[a] χ^2	Unstandardized coefficients X	Y	X^2	XY	Y^2
1. Autonomy	36	25,78**	0,10**	0,00	-0,06***	0,05**	0,02*
2. Integration	75	0,94	0,00	-0,01	0,00	0,00	0,00
3. Involvement	57	9,09	0,09	-0,08	-0,01	0,02	0,02
4. Support	43	3,44	0,01	-0,04	0,01	-0,01	0,01
5. Training	74	19,73*	0,03	-0,03	0,00*	0,01	-0,01*
6. Welfare	59	13,93	0,14*	-0,02	-0,01	0,01	0,02
7. Formalization	70	17,70*	-0,02	-0,07	0,00	0,01	0,01
8. Tradition	36	143,05***	-0,34***	-0,13**	0,04***	-0,03**	0,02**
9. Innovation	53	28,16***	0,06	0,05	-0,01	-0,04**	-0,01
10. Focus	48	41,51***	-0,02	-0,30***	-0,02	-0,07	-0,05
11. Reflexivity	56	6,27	0,06*	-0,03	0,00	-0,01	0,00
12. Clarity	50	7,88	0,17*	0,05	-0,07*	0,02	-0,02
13. Efficiency	73	4,44	-0,04	0,00	0,00	0,01	0,00
14. Effort	68	94,82***	-0,04*	0,04	0,00*	0,01	-0,01
15. Feedback	59	59,29***	0,04	-0,11***	0,00	0,00	0,02***
16. Pressure	58	16,66*	-0,10	0,01	0,04	-0,01	0,01
17. Quality	67	20,21*	0,05	-0,01	-0,03	-0,03*	-0,03**

Note. [a] Probability estimates were corrected for family-wise error using the Holm-Bonferroni method.

* p< .05. ** p< .01. *** p< .001.

Using the robust regression analysis resulted in the model being able to account for 10 of the 17 scales. Tests of the slope and curvature along the line of congruence and the line of incongruence were carried out based on the unstandardized regression coefficients of these 10 scales. The results are presented in table 7.

Table 7

Results from the tests of the line of congruence and the line of incongruence.

		Line of congruence		Line of incongruence	
OCM scale	N	Slope (b_1+b_2)	Curvature ($b_3+b_4+b_5$)	Slope (b_1-b_2)	Curvature ($b_3-b_4+b_5$)
1. Autonomy	36	0,096	0,017	0,099**	-0,083**
5. Training	74	-0,006	-0,008	0,060**	-0,020***
7. Formalization	70	-0,085**	0,019	0,046	-0,004
8. Tradition	36	-0,470***	0,035***	-0,214***	0,104***
9. Innovation	53	0,104	-0,052**	0,014	0,026**
10. Focus	48	-0,318**	-0,152	0,283**	-0,002
14. Effort	68	0,007	0,002	-0,080	-0,020***
15. Feedback	59	-0,073**	0,020	0,156***	0,021
16. Pressure	58	-0,090	0,036	-0,107	0,053
17. Quality	67	0,039	-0,086***	0,065**	-0,025

Note. b_1, b_2, b_3,b_4 and b_5 are the unstandardized regression coefficients on X, Y, X^2, XY and Y^2 from the robust regression analysis (see table 6), respectively.

* p< .05. ** p< .01. *** p< .001.

The test of slopes and curvatures along the lines revealed that 9 of the 10 had either a significant slope or curvature along either the line of congruence or the line of incongruence. As it is beyond the scope of this study to present all 9 response surfaces visually, 3 response surfaces will be presented and their lines linked to the results in table 7.

Figure 2a shows how the relationship between interview statements coded as Strengths and Weaknesses on SWOT on the category Autonomy is related to the predicted scores on the corresponding scale on the OCM questionnaire.

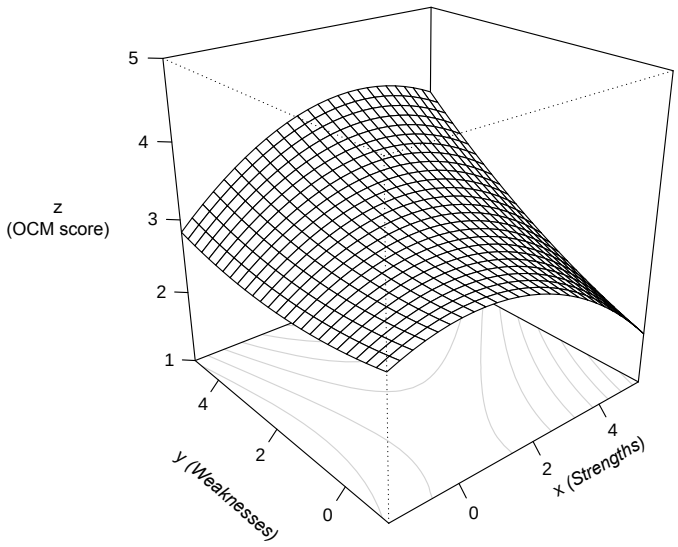

Figure 2a. Surface plot of the relationship between strengths (X) and weaknesses (Y) on the predicted scores on the questionnaire on the scale Autonomy (Z).

A test of the line of congruence, that runs from the near most corner of the surface to the far most corner, shows that neither the slope nor the curvature is significant. Meaning that, a person having both a high frequency of positive statements and a high frequency of negative statements is predicted to have the same score on the questionnaire as a person having a low frequency of both positive and negative statements.

The line of incongruence, running from the left most corner to the right most corner, has a significant negative curvature (-0.083, $p < .01$). This indicates that the scores on the questionnaire is predicted to decrease as the discrepancy between positive and negative statements increase. The line also has a significant positive slope (0,099, $p < .01$), indicating that the associated scores on the questionnaire is predicted to increase as one follows the line of incongruence from the left most corner (where there is an overweight of negative statements) to the right most corner (where there is an overweight of positive statements).

The combination of a significant positive slope and a significant negative curvature along the line of incongruence indicates that scores on the questionnaire is expected to increase as the number of positive statements increase compared to negative statements, but

that this effect tends to diminish when the discrepancy gets very high.

Since the maximum numbers of statements on Strengths (7 statements) is higher than the maximum number of Weaknesses (6 statements), the line of congruence and line of incongruence are a slightly displaced on the response surface presented in figure 2a. The line of congruence is shifted a bit to the left of the far most corner, and the line of incongruence is shifted a bit towards the near most corner along the x-axis. Figure 2b shows a two-dimensional contour plot of the surface with the X-axis trimmed at the axis value of 4,3 (6 statements, which corresponds to the y-axis value of 3 (x statements).

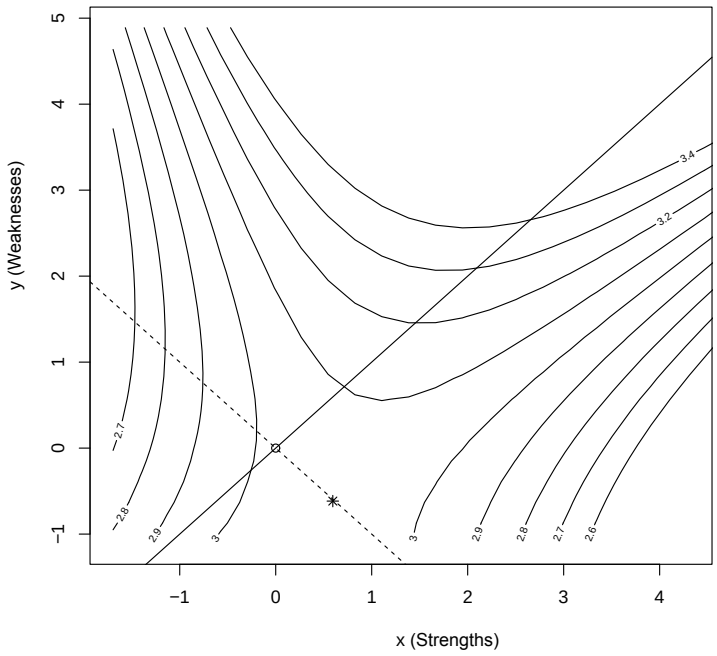

Figure 2b. Contour plot of the surface with the x-axis trimmed at axis value 4,3 (6 statements, corresponding to Y-axis value 6 (6 statements).

The response surface corresponding to the OCM scale Tradition is presented in figure 3a. This surface differs from the other surfaces in that the slope and curvature for both the line of congruence and the line of incongruence are significant.

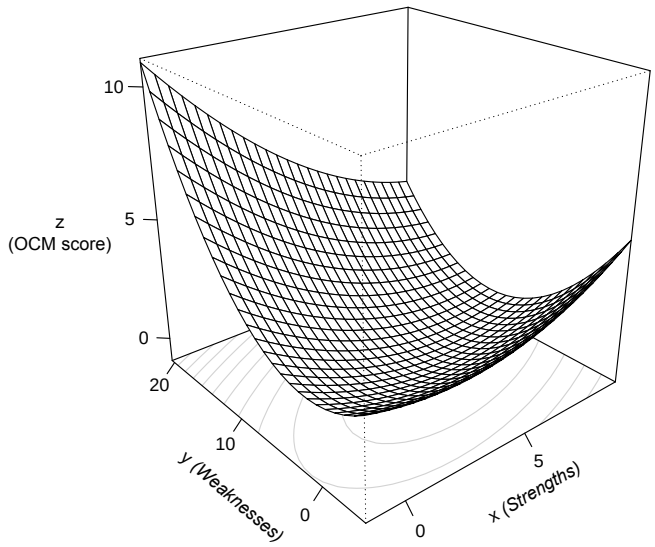

Figure 3a. Surface plot of the relationship between strengths (X) and weaknesses (Y) on the predicted scores on the questionnaire on the scale Tradition (Z).

The line of incongruence has a positive curvature (0,104, $p < .001$) indicating that the aggregated score on the scale is predicted to increase as the discrepancy of positive and negative statements increase. The line also has a negative slope (-0,214, $p < .001$), meaning that persons having an overweight of negative statements are predicted to have a higher score on the scale on the questionnaire than persons having an overweight of positive statements. This trend is only found in the this subset. The line of congruence has a negative slope (-0,470, $p < .001$) combined with a positive slope (0,035, $p < .001$), giving the surface a convex shape along the line. This suggests that the scores on the questionnaire is predicted to decrease following the line from the near most corner, but that there is a change along the line where the predicted scores start to increase. This is a trend along the line of congruence only observed on the scale Tradition. The contour plot of the surface is presented in figure 3b.

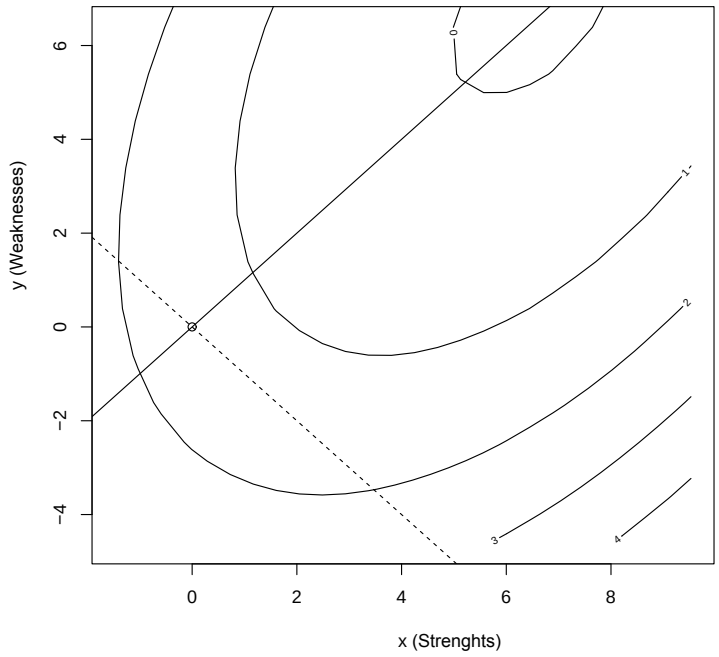

Fig 3b. Contour plot of the surface to Tradition with the Y-axis trimmed at axis value 6,4 (11 statements, corresponding to Y-axis value 9,5 (11 statements)

The response surface corresponding to the OCM scale Innovation is presented in figure 4a. This surface is also unique in this analysis as it is the only surface where the line of congruence and the line of incongruence both have significant curvatures, but neither have significant slopes.

The line of incongruence has a significant positive curvature (0,026, $p < .01$). This indicates that the scores on the questionnaire is predicted to increase as the discrepancy between positive and negative statements increases on either side of the line of congruence. In this case this means that a person having 12 positive statements and 0 negative statements is expected to get the same score on the questionnaire as a person having 0 positive and 12 negative statements. The line of congruence has a significant negative curvature (-0,052, $p < .01$). This means that the score on the questionnaire is predicted to increase as both the degree of positive and negative statements increase. But this trend changes along the line so that the score is predicted to decrease the closer one comes to the far corner of the surface. The combination of a positive curvature along the line of incongruence and a negative curvature along the line of congruence gives the surface distinct saddle shape.

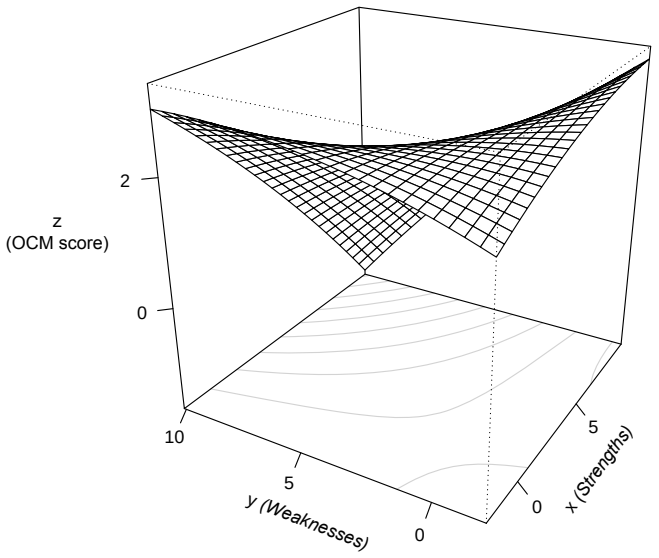

Figure 4a. Surface plot of the relationship between strengths (X) and weaknesses (Y) on the predicted scores on the questionnaire on the scale Innovation (Z).

Figure 4b presents a contour plot of the surface. Since the highest number of statements on both Strengths and Weaknesses is 12, none of the axis have been trimmed.

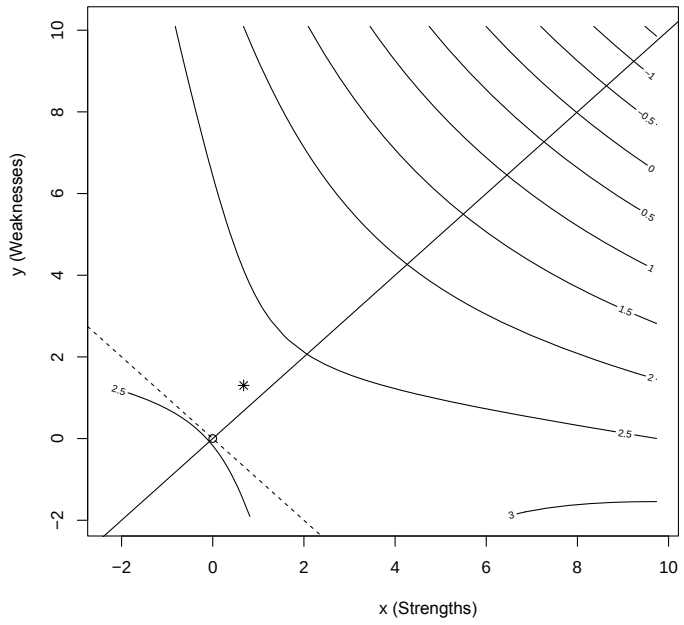

Figure 4b. Contour plot of the surface to Innovation.

Discussion.

Summary of results.

The purpose of this study was to investigate if polynomial regression and response surface methodology could be used to examine the relationship between qualitative data gathered using interviews and quantitative data gather using the OCM questionnaire. The OCM questionnaire consists of 17 scales measuring different aspects of the work environment. An aggregated score was calculated for all participants who completed the survey on each of the 17 scales. Content analysis was used to make the qualitative data from the interviews comparable to the quantitative data from the questionnaires. This involved unitizing the transcripts into a collection of statements. These statements were top-down coded on the four categories of the SWOT and on the 17 categories of the OCM. The content analysis gave all of the participants a sum of the number of statements coded as Strengths and Weaknesses on each of the 17 categories of the OCM. The qualitative and quantitative strands of data were integrated using polynomial regression analysis. Response surface methodology was used to interpret the coefficients from the polynomial regression analysis. A response surface of the predicted values was produced for each of the 17 regression analysis performed. The unstandardized regression coefficients from the regression analysis was used to calculate specific lines of interest across this surface. These lines being the line of congruence and the line of incongruence. These lines were tested to see if they differed significantly from zero. This is the first known attempt to compare information from interviews with quantitative data through the use of polynomial regression and response surface methodology.

General discussion.

The following discussion will focus on how polynomial regression and response surface methodology was used to identify scales where the qualitative and quantitative data converged and diverged. Possible explanations will be given for why the polynomial model failed to explain data on some of the scales. This will be followed by a discussion on inter-rater reliability and the use of robust regression and the presence of outlying values.

Converging data. The benefit of using response surface methodology is that it makes it possible to study how congruence and incongruence between positive and negative statements given during an interview is associated with scores on the different scales of the OCM from the questionnaire. The effect of incongruence between positive and negative statements can be investigated by testing the slope and curvature along the line of

incongruence. If the qualitative data converges with the quantitative data (e.g. that they give the same results), one would assume that a person with more positive statements than negative statements on a category of the OCM would assess the corresponding scale on the OCM questionnaire more favourably compared to a person with more negative statements. Also, it would be reasonable to assume that the more positive statements a person has compared to negative, the more favourably he or she would assess the given scale on the questionnaire. These assumptions are supported by the test of the line of incongruence on 5 of the 10 surfaces in this study. These being Autonomy, Training, Focus, Feedback and Quality.

On the scales Focus, Feedback and Quality, the line of incongruence had a significant positive slope. Meaning that the scores on the corresponding scale on the questionnaire was predicted to increase as one moved along the line from having an overweight of negative statements to having an overweight of positive statements.

On the scales Autonomy and Training, the line of incongruence also had a significant positive slope, but this was combined with a significant negative curvature. This indicates a slightly different form of association between the statements and the scores on the questionnaire. As the discrepancy between number of positive and negative statements increase, the score on the questionnaire is predicted to decrease. But, given the significant positive slope, the rate of decrease is less as the number of positive statements increase from being equal to the number of negative statements than the rate of decrease as the number of negative statements increase from being equal to the number of positive statements. As previously mentioned, this can be seen on the response surfaced by following the line of congruence in a leftward and rightward direction from the point where the line of incongruence crosses the line of congruence. This means that an increasing amount of positive statements compared to negative statements is associated with an increase in the predicted scores on the questionnaire, but after the degree of discrepancy reaches a certain high level this effect levels out. Still, a person having a greater number of positive statements compared to negative statements is predicted to have a higher score on the questionnaire compared to a person with a greater number of negative statements.

Diverging data. While the previously mentioned 5 surfaces indicate a convergence between data from the interview and data from the questionnaire, there are also a couple of surfaces that indicate divergence between the strands of data. One notable surface is the one related to *Innovation*. This surface has a significant positive curvature, but the slope is not significantly different from zero. These results indicate that the predicted value on the

questionnaire increases as the discrepancy between positive and negative statements increase in either direction. The increasing difference between positive and negative statements is associated with higher scores, but whether this difference is positive or negative does not seem to matter. One possible explanation for this surface is the presence social desirability bias on the questionnaire. Participants who had a high degree of negative statements about *Innovation* during the interview still rated *Innovation* favourably on the questionnaire. This is a plausible explanation given the nature of the scale. An example of a statement tapping the scale *Innovation* on the questionnaire is: "assistance in developing new ideas is readily available". Since a majority of the sample consist of participants who exercise some form of leadership, it is conceivable that having to asses these kinds of statements could elicit *a social desirability bias*. Since the interview uses open questions the category Innovation is tapped indirectly, thereby reducing social desirability (Fisher, 1993).

The opposite trend is found on the surface corresponding to *Effort*. Similar to *Innovation*, the line of incongruence lacks a significant slope. But on *Effort* the curvature along the line is significantly negative. Indicating that the associated scores on the questionnaire is predicted to decrease as the discrepancy between positive and negative statements increase in both a positive and negative direction. Again, this could be explained by the presence of a bias on the questionnaire, but this time a tendency of "faking bad". The lower scores on the questionnaire could possibly also be explained in another way. Two the five statements tapping *Effort* on the questionnaire are reversed. And all of the five statements aimed at tapping *Efficiency* are reversed. The *Efficiency* scales is assessed just before the *Effort* scale on the questionnaire. Research has indicated that respondents can develop a cognitive response format while completing questionnaires. When statements are reversed they follow this format and do not notice the difference in wording that reverses scale (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Schmitt & Stults, 1986).

The response surfaces corresponding to Innovation and Effort could also be the result of acquiescence bias (yea-saying and nay-saying), which refers to a tendency in respondents to rate positively worded items more favourably and negatively worded items more unfavourably (Winkler, Kanouse, & Ware, 1982). All of the items tapping the scale Innovation are positively worded and two of five items tapping the scale Effort are negatively worded.

Congruence. Using response surface methodology it is also possible to study how different degrees of congruence between positive and negative statements given during

interviews is related to scores on the OCM questionnaire. Is a person that gives 1 positive and 1 negative statement on a category predicted to have the same score on the questionnaire as a person that gives 10 positive and 10 negative statements? Based on the assumption that degree of incongruence in statements explains differences in predicted scores, one would assume they would get the same score. They appear to be equally positive and negative. Neutral. But, in the example above, it is clear that the person having 10 positive and 10 negative statements has more to say about the given category than a person with only 1 positive and 1 negative statement. By studying the slope and curvature along the line of congruence on the response surface it is possible to predict how different levels of congruence is related to predicted scores. On 6 of the 10 scales the line of congruence had either a significant slope or a significant curvature. This indicates that level of congruence is relevant to understanding how number of positive and negative is related to the scores on the questionnaire. The surfaces of *Formalization, Focus and Feedback* had a line of congruence with a significant negative slope. This indicates that the scores on the questionnaire was predicted to decrease as both the number of positive and negative statements increase. This suggests that participants that gave a lot of both positive and negative statements about the OCM categories, tended to assess the scale on the questionnaire less favourably than participants who had gave fewer statements. On the scales *Innovation* and *Quality* the line of congruence had a significant negative curvature. This means that the associated score on the questionnaire was predicted to increase as the level of both positive and negative statements increase, but only up to a certain level of statements. After this level the scores where predicted to decrease, giving the surface a concave shape along the line of incongruence. Taken as a whole, these results suggests that persons who have a lot of both positive and negative statements about the 6 categories tended to assess the scales more negatively on the questionnaire.

Non-significant scales. While the polynomial model could account for a significant proportion of the variance on 10 of the scales, the model did not fit 7 of the scales. One of these scales was *Integration*. This is the category with the highest number of statements from the top-down coding of the OCM. A possible explanation for the model not fitting this scale might be that the scale is too broad. Integration was defined on the coding scheme as “the extent of interdepartmental trust and cooperation”. The Norwegian police organization is large and complex. Investigators in the police will cooperate with many groups and departments. They will also cooperate across police districts and with specialist branches like

the National Criminal Investigation Service (Kripos). This could lead participants to have many positive statements about the cooperation with some groups and many negative statements about other groups. As participants tend to talk about all groups, it is possible that these statements cancel each other out. It is also possible that participants could be positively biased towards cooperation between departments within their own geographical unit and negatively biased towards cooperation with other geographic units. Leniency biases refers to a tendency in respondents to assess people they know well more favourably than they should (Podsakoff et al., 2003). According to the linguistic intergroup bias model, the language used when describing the behaviour of members of in-groups is different from that used when describing the behaviour of members of out-groups (Maass, 1999; Maass, Salvi, Arcuri, & Semin, 1989).Negative behaviour by members of out-groups tends to be described in abstract terms, while negative behaviour by members of in-groups are described with concrete terms. Incidents of negative cooperation with members of out-groups could be described as something general and typical. Incidents of negative cooperation with members of in-groups could on the other hand be described as a single events that are not representative of how cooperation usually unfolds within the unit. The opposite effect is present in cases of positive cooperation. When talking about intradepartmental cooperation it is therefore plausible that the interviewee could be positively biased towards cooperation within his or her own unit and negatively biased towards cooperation with other units. This could lead the coding of statements on the Integration category to be too broad and make it difficult to compare these to the score on Integration from the questionnaire.

Reflexivity is another OCM scale the model did not fit. Reflexivity was defined in the coding scheme as "a concern with reviewing and reflection upon objectives, strategies, and work processes, in order to adapt to the wider environment". *Reflexivity* can be operationalized both as a cognitive style and as a group process (Schippers, Den Hartog, & Koopman, 2007). During the top-down coding in this phase of the project *Reflexivity* was operationalized as a group process. A statement had to refer to overt group behaviour in order to be coded on the category. Unitizing might be a problem using this definition of reflexivity. Unitizing was done prior to the top-down coding and without any specific model in mind. This could lead the units to be too big to be able to tap this category. An hypothetical unit could be "in this department we have discussed how to improve training". A part of this statement taps the *Reflexivity* category, but the overarching point of the unit is that the department focuses on training. The hypothetical statement would be coded on the category

Training. Since statements cannot be double-coded, the part of the sentence that taps the category *Reflexivity* would not be coded at all. As mentioned, statements had to refer to overt group behaviour in order to be coded on this category. This is an amendment by the author when preparing the coding scheme for the coding of the last 38 interviews. The rationale behind this addition is based on the nature of the interviews being performed. The interviewee is stimulated to reflect by the SWOT framework. A definition of the category reflexivity based cognitive style would qualify most of the participants statements for the category. Also, this would tell us more about the individual than the work environment the individual is a part of. While defining *Reflexivity* as group behaviour may cause the category to be under represented due to prior unitizing, it does make it easier to code statements consistently. It is possible that the students doing the top-down coding last year used a less stringent definition. The number of statements coded on *Reflexivity* could suggest this is the case. A total of 839 statements were coded on the category in total. At a whole, these possible explanations suggest that this category could be very difficult to define in a way that makes top-down coding it on statements consistent.

The polynomial models lacking ability to capture the relationship between statements and scores on the scale *Efficiency* can possibly be explained by how this category was defined in the coding scheme. The category was defined as: "the degree of importance placed on employee efficiency and productivity at work. Being reflexive on the difference between what one is actually doing in relation to the goal of the organization or work group". Many of the statements coded on this category were related to how the interviewee assessed his or her departments ability to achieve goals. The statements reflected how efficient the interviewee considered the department was in meeting these goals. The statements tapping this scale on the questionnaire focused more on how work was organized in order to meet these goals. An example of a statement from the questionnaire is: "productivity could be improved if jobs were organized and planned better". It is possible that the top-down coding on the category tapped something different than the statements used to tap the scale on the questionnaire.

The surface corresponding to the scale *Welfare* was not included among the scales that could be accounted for by the model. Since this study made comparisons between the different surfaces to see if there were any patterns, the Holm–Bonferroni method (Holm, 1979) was used to control for family-wise error. The polynomial model was significant for *Welfare* before correction. This method is considered conservative since it does not take into consideration the correlation between the variables in the different models. Using a more

liberal method could possibly have made a difference. It could also be argued that such a correction was not necessary since the primary aim of the study was to investigate whether response surfaces could be used to analyse the relationship between statements and scores. Meaning that comparison between subsets was only of secondary importance.

Inter-rater reliability. The low inter-rater reliability between coders could also explain why many of the subsets could not be explained by the model. According to Krippendorff (2004) a high degree of agreement is necessary to draw valid interferences from the data. The question is if it is possible to attain a high degree of agreement when coding on as many categories as 17? It is possible that some categories lend themselves to a more reliable coding than others. Categories like Training and Feedback might be coded more reliably than categories like Reflexivity and Involvement, because they are easier to identify. When a person talks about training in a statement, it is easy to spot. The statement should most likely be coded on the category Training. This could explain why the polynomial model seems to extract response surfaces from parts of the data that can be interpreted in a meaningful way.

It is also possible that a large part of the disagreement consists of whether a given statement should be coded as a specific category or if it should be coded as a residual. This could mean that the low agreement between coders lead to a systematic under-coding on some categories. Meaning that one coder might tend to code certain statements as residuals that another coder codes as a specific category. Since the data is aggregated on an individual level, this would not necessarily contaminate the sum of Strengths and Weaknesses of the individuals that are part of the subset, but instead only lead to a lesser number of individuals. This could have happened on the subset for Autonomy. The subset only contains data from 36 participants, and still manages to produce a response surface than is in accordance with how one would expect positive and negative statements to be related to questionnaire scores.

Reflectivity and cognitive dissonance. In the previous paragraphs congruence and incongruence has been interpreted to relate to different degrees of positivity and negativity. A person with a greater number of positive statements compared to number of negative statements is assumed to be more positive. This is an assumption. It is possible that incongruence and incongruence represents something quite different. A person that has a great number of both positive and negative statements could be said to be more reflective on the category. He or she will point out both positive and negative aspects of the area the category is aimed at tapping. This would also mean that persons with a high degree of discrepancy

between positive and negative statements is less reflective on the category. Interpreting the result on the basis of this assumption leads to different conclusions.

Another way of interpreting the effect of increasing discrepancy on survey scores could be to see it as an effect of cognitive dissonance (Festinger, 1957). When a person experiences cognitive dissonance he or she could generate a large number positive or negative statements in order to justify an understanding of the work environment as either positive or negative. This could especially be the case when the interviewee is asked follow-up questions. If the participant feels that his or her opinions on feedback are not entirely consistent, he or she might generate a lot of positive statements in order to justify a previous understanding that the department is good at giving feedback. The interview could be seen as a way of un-freezing the participant by asking him or her to reflect on different aspects of a given area of the work environment. This unfreezing could lead the interviewee to experience cognitive dissonance because the reflections made during the interview could challenge previously held conceptions for the area of work being reflected on. This could lead to a great number of positive statements compared to negative statements if the previous conceptions where positive. And vice versa if the previous conceptions where negative. These alternative views on the nature of congruence and incongruence in statements are of course only speculations. But these alternative explanations could indicate that the difference between positive and negative statements given during interviews not necessarily has to do with how positive or negative a person is towards the aspect of the work the interviewee talks about.

Type I and type II error. In the current study a series of robust regression analyses were used to assess the relationship between statements and questionnaire scores. Wilcox and Keselman (2012) have argued that using least square regression can lead researchers to falsely conclude that there is no association between the variables being tested when the assumptions for doing the analysis are violated. The presence of outlying and influential cases represents such a violation (Tabachnick & Fidell, 2006). Using ordinary least square regression, the polynomial model could only explain a significant portion of variance on 1 of the 17 scales. The robust regression on the other hand, indicated that the model was significant on 10 of the scales. These results indicate that solely relying on ordinary least square regression could lead to a type II error. Meaning that one concludes a relationship does not exist when it really does. It also attests to the importance of investigating outlying and influential cases in the data being used. There are numerous types of robust regression analyses and it is important to choose one that adequately addresses the challenges in the data (Wilcox, 2012). The type of

regression analysis used in this study was selected due to its ability to address the presence of multiple outliers. When conducting as many as 17 analyses and comparing results, it is important to control for family-wise error. As the number of comparisons being made increases, so does the probability of making a type I error. In the current study the Holm-Bonferroni (Holm, 1979) method was used to correct for family-wise error. This is a conservative test and could have led to some scales being wrongly rejected from further analyses. Thereby committing a type II error. The scale Welfare was significant before correction for family-wise error.

Outliers. The presence of outliers is often referred to as a problem. But as Belsey et al (1980) have argued, as long as outliers are not the result of measurement error, these unusual cases can contain some of the most interesting information in the study. By investigating why certain cases are unusual and if there are patterns of outlying cases in the sample, one can achieve a greater understanding of the phenomena being investigated. Since this study utilizes robust regression, the regression weights used in this type of analysis can be compared to more traditional measures of outliers. Thereby following the example of Belsey et al (1980), who compared weights from a similar type of robust regression with more common tests for detecting outliers.

Implications.

Along with the previous studies by Jex et al (1997), Mazzola (2011) and Berglund(2012), this study advocates the use of mixed-methods in comparing data gathered using close-ended and open-ended questions. As such, it shows that the combination of qualitative and quantitative data within the same research design can contribute to a more nuanced knowledge than data based on one approach in isolation.

The aim of this study was to see if polynomial regression and response surface methodology could be used to assess the relationship between qualitative data gathered from interviews and quantitative data gathered from questionnaires. The results indicate that this is possible, as the polynomial model could account for the relationship between the two strands of data on 10 of 17 scales. The majority of the response surfaces could also be interpreted in a meaningful way. This technique could therefore contribute to the growing repertoire of methods within the field of mixed-methods, as it integrates the qualitative and quantitative strands of data in a new way.

An analysis of the line of congruence and the line of incongruence on the response surfaces showed that the qualitative and quantitative strands of data converged on 5 scales.

Along with the surfaces that showed diverging results, this indicates that the technique can be used as an instrument to validate quantitative questionnaires.

The results from the study also predicted that people with congruent numbers of positive and negative statements were associated with different scores on the OCM scales depending on the number of statements they gave. This casts light on a group of participants that would be overlooked using techniques based on difference scores.

The presence of significant curvatures along the surface also indicate that the relationship between statements given during interviews and scores on the questionnaire scales could be complicated. This suggests that the use of difference scores alone might not be capable of assessing the relationship.

The use of this technique in psychology has so far primarily been used to study how differences between self and others relate to an outcome within multisource feedback research (Shanock et al., 2010). As a result, its use has primarily been based on independent variables measured using self-report questionnaires with Likert-scales. The results of this study indicates that the methodology can be used within other fields of psychology and with other types of measures.

Limitations.

This study has five limitations. The low inter-rater reliability between coders on the top-down coding on the OCM scales is a notable limitation. Krippendorff (2004) argues that a high level of reliability is necessary in order to make valid inferences. One of the drawbacks of using interviews is that the content analysis of the resulting transcripts is time consuming. While content analysis encourages an iterative approach to the coding on categories, this is not always possible within a set timeframe. The coding of the statements across two separate phases by different coders, could also have contributed to low reliability. The existence of tacit rules for coding is a possibility. This could have led to slightly different coding on some of the scales. A process that involved a top-down coding of all the statements on the OCM over several iterations could possibly have improved agreement.

The top-down coding was done in two stages. In the first stage all the statements were coded on the SWOT and IGLO, and in the second stage all the statements were coded on the scales of the OCM. This could have influenced the way positive and negative statements were distributed on OCM categories. For example, statements coded as positive during the first stage, could represent something negative when coded on the OCM.

A third limitation is related to the sample size. On some of the subsets used in the

regression analyses the number of participants was small. This could have had an impact on the results, as the effects being tested could have been too weak to be captured with such a small sample. Using a larger sample could have yielded different patterns across the response surfaces.

A fourth limitation regards the technique used to compare the two samples. By using other approaches, such as variance analysis, the results could have been different. The use of polynomial regression and response surface methodology using the raw frequency of statements aggregated on an individual level, is of course experimental.

A fifth limitation regards the use of follow-up questions during interviews. These questions could have generated a larger number of statements on some of the scales being examined in this study. However, as all of these questions were asked after the main questions, it is possible to remove these statements in future analyses.

Future studies.

In this study the raw frequency of positive and negative statements on an individual level was used as independent variables. This makes the response surfaces difficult to interpret. An alternative approach could be to assess each participants degree of positivity and negativity on a scale with fixed minimum and maximum values. For example, by giving each participant a score ranging from 1 to 3 on positivity regarding a given scale. Where 1 indicated that the participant was not notably positive and 3 that he or she was very positive. By doing the same with regard to how negative each participant was, one could use these as independent variables in a polynomial regression analysis. This would make the surfaces easier to interpret as these scales could be centred at their midpoint. This could also make it easier to draw inferences from the results, as the scales would represent degree of positivity and negativity. As mentioned previously, the number of positive and negative statements do not necessarily represent how positive or negative a participant is with regard to the scale in question.

This study used statements coded on all levels of IGLO. A future study could investigate whether a matching of statements to specific levels of the organization provides different patterns on the response surface.

The 17 scales that constitute the OCM are distributed on the 4 quadrants of the competing values model. It might not be possible to achieve a high level of agreement between coders when using this many categories. One potential solution could be to code each statement on a scale from each of the quadrants, thereby reducing insecurity in the coder about which scale the statement belonged to. It is possible that one statement could represent

more than one scale. By coding each statement on categories from each of the quadrants, the coder would be given the opportunity to code the same statements on several categories as long as these belonged to different quadrants. Another approach could be to only code the statements on a selection of the scales instead of utilizing the whole model.

Conclusion

The aim of this study was to investigate if polynomial regression analysis and response surface methodology could be used to examine the relationship between positive and negative statements from interviews and scores on the OCM questionnaire. The main finding was that this is possible. Results indicated that response surfaces could be used to examine both convergence and divergence between the qualitative and quantitative strands of data. This study also showed that this technique could provide a more nuances picture of this relationship than alternative approaches using difference scores.

References

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology, 22*(3), 309-328.

Belsey, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*: John Wiley.

Berglund, L. M. (2012). *The use of mixed-methods when analysing work components in police investigation.* (Master), University of Oslo, Oslo.

Bollen, K. A., & Jackman, R. W. (1990). Regression diagnostics: An expository treatment of outliers and influential cases. In J. Fox & J. S. Long (Eds.), *Modern methods of data analysis* (pp. 257-291). Newbury Park, CA: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. doi: 10.1037/h0046016

Cohen, A., Nahum-Shani, I., & Doveh, E. (2010). Further insight and additional inference methods for polynomial regression applied to the analysis of congruence. *Multivariate Behavioral Research, 45*(5), 828-852. doi: 10.1080/00273171.2010.519272

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 15-18.

Creswell, J. W. (2011). Controversies in mixed methods research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 269-283). Thousand Oaks, CA: Sage Publications.

Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes, 64*(3), 307-324. doi: http://dx.doi.org/10.1006/obhd.1995.1108

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*(3), 265-287.

Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *The Academy of Management Journal, 36*(6), 1577-1613. doi: 10.2307/256822

Edwards, J. R., & Rothbard, N. P. (1999). Work and family stress and well-being: An examination of person-environment fit in the work and family domains. *Organizational Behavior and Human Decision Processes, 77*(2), 85-129.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, Calif.: Stanford University Press.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*(2), 303-315. doi: 10.2307/2489277

Freelon, D. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science, 5*(1), 20-33.

Hoaglin, D. C., & Welsch, R. E. (1978). The hat matrix in regression and ANOVA. *The American Statistician, 32*(1), 17-22.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70. doi: 10.2307/4615733

Jex, S. M., Adams, G. A., Elacqua, T. C., & Lux, D. J. (1997). A comparison of incident-based and scale measures of work stressors. *Work & Stress, 11*(3), 229-238. doi: 10.1080/02678379708256837

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133. doi: 10.1177/1558689806298224

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks, Calif.: Sage.

Lund, T. (2011). Combining Qualitative and Quantitative Approaches: Some Arguments for Mixed Methods Research. *Scandinavian Journal of Educational Research, 56*(2), 155-165. doi: 10.1080/00313831.2011.568674

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. In P. Z. Mark (Ed.), *Advances in Experimental Social Psychology* (Vol. Volume 31, pp. 79-121): Academic Press.

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology, 57*(6), 981-993. doi: 10.1037/0022-3514.57.6.981

Mazzola, J. J., Schonfeld, I. S., & Spector, P. E. (2011). What qualitative research has taught us about occupational stress. *Stress and Health, 27*(2), 93-110. doi: 10.1002/smi.1386

Milne, B., & Powell, M. (2010). Investigative interviewing. *The Cambridge handbook of forensic psychology*, 208.

Myers, R. H., & Anderson-Cook, C. M. (2009). *Response surface methodology: process and product optimization using designed experiments* (Vol. 705): John Wiley & Sons.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Patterson, M. G., West, M. A., Shackleton, V. J., Dawson, J. F., Lawthom, R., Maitlis, S., . . . Wallace, A. M. (2005). Validating the organizational climate measure: Links to managerial practices, productivity and innovation. *Journal of Organizational Behavior, 26*(4), 379-408. doi: 10.1002/job.312

Plano Clark, V. L. (2010). The adoption and practice of mixed methods: U.S. trends in federally funded health-related research. *Qualitative Inquiry, 16*(6), 428-440. doi: 10.1177/1077800410364609

Ployhart, R. E., Schmitt, N., & Schneider, B. (2006). *Staffing organizations: contemporary practice and theory*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. doi: 10.1037/0021-9010.88.5.879

Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science, 29*(3), 363-377. doi: 10.2307/2631061

Rousseeuw, P., & Yohai, V. (1984). Robust regression by means of S-estimators *Robust and nonlinear time series analysis* (pp. 256-272): Springer.

Salibian-Barrera, M., & Yohai, V. J. (2006). A fast algorithm for S-regression estimates. *Journal of Computational and Graphical Statistics, 15*(2).

Schippers, M. C., Den Hartog, D. N., & Koopman, P. L. (2007). Reflexivity in teams: A measure and correlates. *Applied Psychology, 56*(2), 189-211. doi: 10.1111/j.1464-0597.2006.00250.x

Schmitt, N., & Stults, D. M. (1986). Methodology Review: Analysis of Multitrait-Multimethod Matrices. *Applied Psychological Measurement, 10*(1), 1-22. doi: 10.1177/014662168601000101

Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational climate and culture. *Annual Review of Psychology, 64*(1), 361-388. doi: doi:10.1146/annurev-psych-113011-143809

Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for

examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology, 25*(4), 543-554.

Skinner, B. F. (1957). *Verbal behavior*: Appleton-Century-Crofts New York.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics*. Boston: Pearson/Allyn and Bacon.

Tashakkori, A., & Teddlie, C. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 3-50). London: Sage.

Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 1-41). Thousand Oaks, CA: Sage Publications.

Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing*. Amsterdam: Elsevier/Academic Press.

Wilcox, R. R., & Keselman, H. J. (2012). Modern regression methods that can substantially increase power and provide a more accurate understanding of associations. *European Journal of Personality, 26*(3), 165-174.

Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology, 67*(5), 555.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, 642-656.

Appendix A - The SWOT Interview Guide

1.

Please tell us what you think works well regarding the investigative work here in the police district - we call this the strength of the investigative work.

2.

Please tell us what you think does not work well regarding the investigative work here in the police district – we call this the weakness of the investigative work.

3.

Please tell us what you consider to be opportunities for improving the investigative work here in the police district – we call this the opportunities in the investigative work.

4.

Please tell us what you consider to be threats against improving the investigative quality here in the police district - we call this the threats in the investigative work.

Appendix B - OCM Questionnaire

Undersøkelse av organisasjonsklimaet i Politiet

Prosjektet «Kvalitet i etterforskningen»

Spørsmål	Helt Feil	Ganske feil	Ganske riktig	Helt riktig
1. Ledelsen lar stort sett ansatte ta sine egne beslutninger				
2. Ledelsen har tillit til at folk kan ta arbeidsrelaterte beslutninger uten å innhente tillatelse først				
3. Ledelsen holder streng kontroll med arbeidet til de ansatte				
4. Ledelsen har for strengt regime over måten ting blir gjort på				
5. Det er viktig å dobbeltsjekke med nærmeste leder før man tar en beslutning				
6. Folk er mistenksomme overfor andre avdelinger				
7. Det er svært lite konflikt mellom avdelingene her				
8. Folk er innstilt på å dele informasjon på tvers av avdelinger				
9. Det er svært effektivt samarbeid mellom avdelingene				
10. Det er lite respekt mellom noen av avdelingene her				
11. Ledelsen lar de ansatte medvirke i beslutninger som angår dem				
12. Endringer blir gjort uten å snakke med de involverte				
13. Folk har ingen innvirkning i avgjørelser som påvirker arbeidet deres				
14. Folk føler at beslutninger ofte tas uten at de blir hørt				
15. Informasjon deles i stor grad				
16. Det er ofte kommunikasjonssvikt her				
17. Overordnede er svært dyktige til å forstå folks problemer				
18. Overordnede viser at de har tiltro til sine ansatte				
19. Overordnede hos oss er vennlige og lette å henvende seg til				

20. Folk kan stole på at overordnede gir god veiledning				
21. Overordnede viser forståelse for sine ansatte				
22. Folk får ikke tilstrekkelig opplæring i nye systemer eller nytt utstyr				
23. Folk får tilstrekkelig opplæring i bruk av nye systemer og utstyr				
24. Organisasjonen gir kun et minimum av den opplæringen folk trenger for å gjøre jobben sin				
25. Folk blir sterkt oppmuntret til å utvikle sine ferdigheter				
26. Denne organisasjonen vier lite oppmerksomhet til ansattes interesser				
27. Denne organisasjonen forsøker å ta vare på sine ansatte				
28. Denne organisasjonen bryr seg om sine ansatte				
29. Denne organisasjonen prøver å handle rettferdig overfor sine ansatte				
30. Hos oss blir det oppfattet som svært viktig å følge reglene				
31. Folk kan ignorere formelle prosedyrer og regler hvis det bidrar til å få jobben gjort				
32. Hos oss må alt gjøres etter reglene				
33. Hos oss er det ikke nødvendig å følge alle prosedyrer til punkt og prikke				
34. Hos oss blir ingen særlig opprørt hvis reglene brytes				
35. Toppledelsen foretrekker å holde seg til de etablerte, tradisjonelle måtene å gjøre ting på				
36. Måten denne organisasjonen gjør ting på har aldri forandret seg særlig mye				
37. Ledelsen er ikke interessert i å prøve ut nye ideer				
38. Hos oss skjer endringer i måten ting gjøres på svært langsomt				
39. Hos oss blir nye ideer gjerne akseptert				
40. Organisasjonen reagerer raskt når endringer er nødvendig				
41. Behov for å gjøre ting annerledes fanges raskt opp av ledelsen				
42. Denne organisasjonen er svært fleksibel; den kan raskt endre prosedyrer for å møte nye vilkår, og problemer løses når de oppstår				

43. Støtte til utvikling av nye ideer er lett tilgjengelig				
44. Folk i denne organisasjonen er alltid ute etter å se problemer fra nye vinkler				
45. Denne organisasjonen er ganske innadrettet; man bryr seg ikke om hva som skjer i samfunnet				
46. Det legges ikke mye vekt på måter å bedre service til publikum				
47. Publikum sitt behov er ikke ansett som topp prioritet hos oss				
48. Denne organisasjonen er treg til å reagere på publikums behov				
49. Denne organisasjonen ser stadig etter nye muligheter i samfunnet				
50. Måten de ansatte jobber sammen på i denne organisasjonen endres gjerne hvis det bedrer prestasjonen				
51. Arbeidsmetodene brukt i denne organisasjonen blir ofte diskutert				
52. Hvorvidt de ansatte jobber effektivt sammen, blir regelmessig diskutert				
53. Denne organisasjonens målsetninger endres i takt med forandringer i miljøet				
54. I denne organisasjonen tar man seg tid til å evaluere organisasjonens målsetninger				
55. Ansatte har en god forståelse av organisasjonens formål				
56. Organisasjonen sin fremtidige retning blir klart og tydelig kommunisert til alle				
57. Ansatte har ikke en klar forståelse av hva som er organisasjonens mål				
58. Alle som jobber her er bevisst på organisasjonens fremtidsplaner og retning				
59. Det finnes en klar oppfatning av hvor organisasjonen går				
60. Tid og penger kunne blitt spart dersom arbeidet var bedre organisert				
61. Ting kunne blitt gjort mye mer effektivt hvis folk tok seg tid til å tenke seg om				
62. Dårlig planlegging resulterer ofte i at man ikke når sine målsetninger				
63. Produktiviteten kunne blitt forbedret om arbeidet ble bedre organisert og planlagt				
64. Hos oss ønsker folk alltid å prestere så godt de kan				
65. Folk er entusiastiske i forhold til jobben sin				

66. Her slipper folk unna med å gjøre så lite som mulig				
67. Folk er innstilt på å gjøre en ekstra innsats for å utføre en god jobb				
68. Her legger ikke folk mer innsats i arbeidet sitt enn det de må				
69. Folk får som regel tilbakemelding i forhold til kvaliteten på det arbeidet de gjør				
70. Folk har ingen anelse om hvorvidt de gjør en god jobb				
71. Det er generelt vanskelig for ansatte å vurdere kvaliteten på det de presterer				
72. Folks prestasjoner måles regelmessig				
73. Måten folk gjør jobben sin på blir sjelden evaluert				
74. Det forventes for mye av folk i løpet av en dag				
75. Vanligvis er ikke folks arbeidsbelastning spesielt krevende				
76. Ledelsen krever at folk jobber ekstremt hardt				
77. Folk er under sterkt press for å nå målsetninger				
78. Arbeidstempoet her er ganske avslappet				
79. Denne organisasjonen forsøker alltid å oppnå de høyeste kvalitetsstandardene				
80. Hos oss blir kvalitet tatt svært seriøst				
81. Folks oppfatning er at organisasjonens suksess avhenger av høy kvalitet på arbeidet				
82. Denne organisasjonen har ikke noe særlig rykte for å levere tjenester av topp kvalitet				

Appendix C - Coding Scheme SWOT/IGLO

Hva er et meningsbærende utsagn (statement)? – Basert på prosedyre for kull 2011

Introduksjon:

- I innholdsanalysen kan man beskrive en enhet som et identifiserbart budskap eller en budskapskomponent. (Neuendorf s 71)
- Enheter kan være ord, karakterer, tema, tidsperioder, samhandlinger eller andre resultater av å "dele opp 'kommunikasjon' i biter. (Neuendorf s 71)
- Generelt kan man si at enheter er helheter som den som analyserer vil skille mellom og behandle som uavhengige elementer. For eksempel dersom man skal telle, må objektene som telles være distinkte både konseptuelt og logisk, kanskje også fysisk, ellers vil ikke det numeriske utfallet gi noen mening. Å telle meninger er problematisk dersom det er ikke er mulig å skille mellom meninger og sikre seg at en mening ikke er avhengig av en annen. (Klippendorf s. 97)
- Det er ønskelig å definere beskrivende enheter som de minste enhetene som inneholder all den informasjonen man trenger til analysen. (Klippendorf s. 100)
- Oppfølgingsspørsmål settes i klamme [oppfølgingsspørsmål] foran utsagnet, for å klargjøre hva innholdet/utsagnet handler om.
- Skrive poenget med uniten i kommentarfeltet for å klargjøre hva fokuset i setningen er.

Definisjon:

- Et utsagn er definert som den minste meningsbærende enhet som uttrykker et logisk, konsistent og separat synspunkt. Et utsagn kan være del av en setning, det kan være en hel setning eller flere setninger. . *(Hoff et. al, 2009 s 14)*
- *Stort nok til å være meningsfullt og samtidig formidle det informanten ønsker å formidle.*

Hvordan definere SWOT-R- kategoriene? Basert på prosedyre for kull 2011

- Utsagn som er hentet fra transkripsjonene kodes i SWOT-R format. Dette vil si i kategoriene styrker, svakheter, muligheter eller trusler, samt residualer.

SWOT-R-kategoriene

- Styrker og svakheter defineres som utsagn som rettes mot her-og-nå status i organisasjonen, mens muligheter og trusler defineres som utsagn som er rettet mot fremtid og eksternt til organisasjonen.
- SWOT-delen av M-SWOT går altså ut på at man tar hvert utsagn i et intervju og kategoriserer det som en styrke, svakhet, mulighet eller trussel. De fire kategoriene bestemmes utfra tre dimensjoner: positiv/negativ, nåtid/fremtid og intern/ekstern.
- Når det gjelder dimensjonene intern og ekstern, setter vi spørsmålstegn om hvorvidt disse utelukkende kan gjelde kun noen av kategoriene. Vanligvis vil man definere styrker og svakheter i forhold til interne forhold, og muligheter og trusler opp mot eksterne forhold. Her mener vi at disse kan gli over i hverandre, men at tidsperspektivet styrer da kodingen.
- Ved vanskeligheter: Vi må først og fremst kode ut ifra konteksten utsagnet fremkommer. Dersom vi ikke klarer å kategorisere enheten med hjelp av kontekst, kan man se om det hjelper å kode ut fra hvilket spørsmål informanten svarer på. "Vi har mange opplæringsdager" kan være alle fire kategoriene, samt residual, avhengig av hvilket spørsmål som blir stilt.

Residual

- Et nøytralt utsagn som ikke faller inn under SWOT kategoriene. I tillegg vil kontekstuelle utsagn (som for eksempel at noen kommer inn døren og spør om man vil ha kaffe) også falle utenfor.

Eksempler:

***Styrker*: Positiv nåtid**

”vi har konkurranse om hver eneste stilling, og det gjør at vi vi greier og og tilegne oss ehh høy kompetanse”

***Svakheter:* Negativ nåtid**

”det som er helt åpenbart som ligger i bunn det er at, det er at antall saker som kommer inn eh som vi i hvert fall burde ha sett på og ha vært å etterforsket, det er langt større enn det vi makter”

***Muligheter:* Positiv fremtid**

”større fleksibilitet ja det vil føre oss føre oss lenger fram”

***Trusler*: Negativ fremtid**

”det som vi opplever eh her på huset, det er eh det er noe så enkelt som faren for at lønnsnivået på operativt asså det å gå i blå skjorte, at det blir annerledes og bedre, eh mer lukrativt enn det å gå på etterforskning”

***Residual:* nøytralt utsagn**

”Det det vet dere jo”

Kakeoppskrift:

- 1.steg: bryter ned i så korte meningsbærende enheter som mulig uten å kategorisere i SWOT. (for å unngå forutinntatthet).
- 2.steg: kode i SWOTr, altså med residualer.
- Dersom vi ikke klarer å plassere et utsagn inn i SWOT-kategoriene etter å ha satt det inn i kontekst og/eller forstått det ut ifra spørsmålet forut for utsagnet, kategoriserer vi det som en residual. Eks: Må bare drikke litt her. Helt nøytralt uttrykk.
- Fremtid vs. nåtid:
    - Normativ = fremtid: setninger som sier noe om hvordan noe kan være, bør være, skal, vil, må være. Ordet "viktig" signaliserer gjerne en mulighet.
    - Deskriptiv = nåtid: setninger som sier noe om hvordan ting er nå, per i dag.

Hvordan definere IGLO-kategoriene - Basert på prosedyre for kull 2011

- **Individ-nivå:** Refererer til individets opplevelser, meninger og følelser. Hvordan en personlig opplever endringer, arbeidsform, stress, krav, attribuere til enkeltperson etc.
- **Gruppe-nivå:** Refererer til avdelingen, seksjonen, grupper og team. Vi-pronomen, Arbeidsmiljøet her hos oss. Samarbeid her hos oss. Mellommenneskelige forhold. Politistasjon/lensmannskontor. Organisering av interaksjon på gruppenivå.
- **Ledelses-nivå:** referer til nærmeste leder, ulike lederadferder og generell ledelse
- **Organisasjonsnivå:** refererer til organisasjonsadferd på flere steder, aktiviteter og holdninger knyttet til organiseringen, strukturer og strategier, organisatoriske visjoner og mål, og verdier. Distriktet som helhet. Samspill mellom tjenestesteder. Organisering av interaksjon på gruppenivå med øvre føringer. POD (politidirektoratet)
- **Ekstern:** 5. Eksterne overordnede pålegg/føringer som dokumenter, rundskriv, styringsdokumenter etc. Riksadvokaten
- **Samfunn:** 6. Beskrivelser av samfunnsbildet, publikum og kriminalitetsbildet. Vi har mange voldssaker. Ikke når det er i forhold til noe som har med politiet å gjøre. Det som ikke hører til noen andre steder enn residual.
- **Residual:** 7. Det som ikke hører til noen andre organisasjonsnivåer. Fyll som "jeg har ikke mer å si noe om det" er residual (både på iglo og swot), og ikke individnivå. Står alene, er nøytralt og beskrivende.

Kakeoppskrift:

- På hvilket organisasjonsnivå ligger svakheten på? Tenk hvor ligger svakheten, hvor er det problemet ligger? Hvor ligger styrken, svakheten, mulighetene og truslene?
- Politimesterens "vi" er på distriktsnivå og derfor org, med mindre det presiseres at det er på et konkret sted, som Alta.

    - Tatt utgangspunkt i Ellen Flakkes definisjoner fra 2008, og det vi definerte på møtet den 04.11.2011, og møtet 09.10.2012

Appendix D - Coding Scheme OCM

1. The human relations approach.

internal focus and flexibility in relation to the environment.

reflects the tradition derived from the socio-technical and human relations. This approach emphasizes the well-being, growth and commitment of the community of workers within an organization.

The Human Relations has norms and values associated with belonging, trust, and cohesion, achieved through means such as training and human resource development. Coordination and control are accomplished through empowerment and participation, and interpersonal relations are supportive, cooperative, and trusting in nature.

1. autonomy - designing jobs in ways which give employees wide scope to enact work

2. integration - the extent of interdepartmental trust and cooperation

3. involvement (participation and communication)

a) participation - employees have considerable influence over decision-making

b) communication - the free sharing of information throughout the organization

4. supervisory support - the extent to which employees experience support and understanding from their immediate supervisor

5. emphasis on training - a concern with developing employee skills

6. employee welfare - the extent to which the organization values and cares for employees

2. The internal process approach.

internal focus and tight control within the organization.

reflects a Tayloristic concern with formalization and internal control of the system in order that resources are efficiently used.

Emphasis is on stability, where the effects of environmental uncertainty are ignored or minimized. Coordination and control are achieved by adherence to formal rules and procedures. The Internal Process Model represents the classic bureaucracy.

7. formalization - a concern with formal rules and procedures

8. tradition- the extent to which established ways of doing things are valued

3. The open systems approach.

external focus and flexible relationships with the environment.

Emphasizes the interaction and adaptation of the organization in its environment, with managers seeking resources and innovating in response to environmental (or market) demands.

9. innovation and flexibility

a) flexibility: an orientation toward change

b) innovation: the extent of encouragement and support for new ideas and innovative approaches

10. outward focus - the extent to which the organization is responsive to the needs of the customer and the marketplace in general

11. reflexivity - a concern with reviewing and reflecting upon objectives, strategies, and work processes, in order to adapt to the wider environment

4. The rational goal approach.

external focus but with tight control within the organization.

Reflects a rational economic model of organizational functioning in which the emphasis is upon productivity and goal achievement.

Primary emphasis is on the pursuit and attainment of well-defined objectives, where norms and values are associated with productivity, efficiency, goal fulfillment, and performance feedback.

12- clarity of organizational goals - a concern with clearly defining the goals of the organization

13. efficiency - the degree of importance placed on employee efficiency and productivity at work. being reflexive on the difference between what one is actually doing in relation to the goal of the organization or work group

14. effort - how hard people in organizations work towards achieving goals. how hard people try to achieve goals.

15. performance feedback - the measurement and feedback of job performance

16. pressure to produce - the extent of pressure for employees to meet targets

17. quality - the emphasis given to quality procedures.